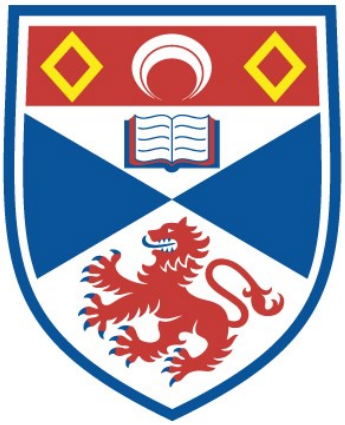
SONG CHARACTERISTICS AND SEXUAL SELECTION IN THE WILLOW WARBLER ('PHYLLOSCOPUS TROCHILUS')

Diego Gil

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1998

L

SONG CHARACTERISTICS AND SEXUAL SELECTION IN THE WILLOW WARBLER (*Phylloscopus trochilus*)

Diego Gil

Thesis submitted in application for the degree of Doctor in Philosophy (Ph.D.) to the University of St Andrews

ABSTRACT

In many species of birds, males have complex song repertoires, which are used in the context of breeding. The succinct rationale of the research done in this thesis is the following paradox: assuming that song repertoires are costly to produce and store, why do males have repertoires of songs rather than a single song? A possible reason is that, if only males of good quality or in good condition are able to produce these repertoires, good quality males would outweigh the costs of the repertoire by increased benefits in mating success. I studied this issue in a population of willow warblers (*Phylloscopus trochilus*). I found some evidence that repertoire size was correlated with male quality. Male age was found to correlate with repertoire size, and there was a positive correlation between repertoire size and survival. The probability of a male having an offspring recruiting into the population was also correlated with repertoire size. Number of fledglings was positively correlated with repertoire size, even when the effect of arrival date was taken into account. However, female choice bore no relation to repertoire size. Females paired first with early arriving males. Females did not use repertoire size either when choosing extra-pair partners. Another characteristic, song length, seemed to be behind this choice, and males with short songs were more likely to be cuckolded. The effects of song repertoires in male-male competition were examined by means of a playback experiment. The results did not provide conclusive evidence of an effect of repertoire in male-male competition. Taken together, the evidence gathered in this thesis suggests that, although repertoire size correlates with several measures of male quality in this species, the maintenance of this trait does not seem to be based on a disproportionally higher mating benefit.

Declarations

(i) I, Diego Gil, hereby certify that this thesis, which is approximately 46.000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a high degree.

date 24.7.98 signature of candidate

(ii) I was admitted as a research student in October 1993 and as a candidate for the degree of PhD in October 1993; the higher study for which this is a record was carried out in the University of St Andrews between 1993 and 1998.

date 24.7.98 signature of candidate

(iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date 24th July 1998 signature of supervisor

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

date 24.7.98 signature of candidate

A MIS PADRES

ACKNOWLEDGEMENTS

I am extremely grateful to Peter Slater, for welcoming me in his group and for his constant encouragement. He was always there to answer my questions, and was a friend rather than a boss. Thanks also to Elizabeth Slater for her wonderful hospitality and cuisine, and for lending me her car when mine broke down in Kippo. Jim Cobb should not be mentioned in this page, but in the front one as the co-author of the thesis. He ringed most of the birds, found most of the nests, and was always happy to discuss the inalienable bloody-mindedness of our dear warblers. More importantly, he and Calla became my Maecenates, provided me with vegetables and fed me delicious dishes. My successful naturalisation in Scotland was possible thanks to my best friend in all these years and more to come, Ihab. I am specially indebted to him for painstakingly teaching me the most colourful swearing in Cairine Arabic, and introducing me to Um Khaltoum. Yannis did his best as far as Greek swearing was concerned, and was always a good Mediterranean to be with. Katharina and Vincent were the best mates one could have hoped for. A million thanks to the E-floor bunch, for being always helpful, and ready for a party or a pint: Patrice, Paul, Simon, Adrian, Rob, Susanne, Alex, Fiona, Richard, Sîan, Iain, Jon, Charles, Dawn, Ronan, Sabine, Jenny, Liz, Leon, et al. Jeff Graves taught me from how to use a Gilson pipette to automatic sequencing. The E-floor is truly inconceivable without his yelling along the corridor. Mike Ritchie put up with the most absurd/daft questions on stats and genetics and introduced me to the mysteries of the cubic splines (still a mystery to me). Nigel Mann helped me with observations in 1996 and was the first human to witness a EPC in willow warblers. Constantino arrived late but has become an essential colleague with whom to talk about women, football and science. Daniel Margoliash kindly sent me a program written by him to calculate Levenshtein distances. Thanks to Richard Griffiths for sending me the sex-linked primers, aka 'the Griffiths sex-machine'. Jeff Podos happily corresponded with me about the exciting dilemmas of using moats to cut trees (he knows what I mean). Benjamín Pérez Casado spent the whole Christmas holiday of 1994 writing the terrific tailor made C-program to analyse willow warbler song sequences, and pretended to believe my assurances about the highly commercial value of such a package. I am indebted to the several grant giving bodies which funded my research throughout these years: the La Caixa/British Council higher degree program, the Russell Trust of the University of St. Andrews, and the TMR fellowship program of the European Community. Thanks to my parents for all their funding, and the encouragement that have they always given me. Hélène es lo mejor que me ha ocurrido en todo este tiempo, soportó con infinita paciencia mis madrugadas pajariles y nunca ha dejado de apoyarme. Merci.

Die heidnischen Stimmen der Vögel.

ELIAS CANETTI

TABLE OF CONTENTS

GENERAL INTRODUCTION

In most bird species of the temperate regions, only males sing (Catchpole & Slater 1995). Bird song mostly occurs in the context of territory defence and mate attraction. This functional link has been demonstrated by experiments in the field, which have shown that song helps to keep out intruders from territories (e.g. McDonald 1989), and to attract females (e.g. Eriksson & Wallin 1986). In many bird species, males have several different copies of their species-specific songs which are used in the same context, and that seem to be equivalent in their message (e.g. Slater 1981; Todt & Hultsch 1996). Bird song shares with other secondary sexual traits an elaborated design that seems to exceed the needs of its function. This redundancy, parallel to the ornamentation of a peacock's tail, suggests that the evolution of song repertoires has been driven by sexual selection (Darwin 1871). A further parallel can be drawn between the two main functions of song, mate attraction and territory defence, and the two modes in which sexual selection acts, female choice and male-male competition.

A demonstration that a trait is being sexually selected should show that individual variation in the expression of the trait correlates with variation in mating success. The mechanism by which differential mating success should come about can either be a female preference to mate with individuals which present an extreme value in the trait, or an advantage at competing with other males for breeding resources, like mates or territories. Several experiments have shown that repertoires are more effective at avoiding take-over of territories by other birds than single songs (Krebs et al. 1978; Yasukawa 1981). The evidence for mate choice is mixed, and whereas most lab studies show that females have a higher sexual response to repertoires than to single songs, many field studies have failed to find any significant effect (Searcy 1992). This pattern has led to the suggestion that song repertoires may be used by females as a cue to male quality in the context of extra-pair paternity (Searcy 1992).

Current models of sexual selection can be divided in two main groups (Andersson 1994). A first group, the Fisherian or arbitrary trait model, proposes that female preference for male traits is maintained by a process of genetic coupling between preference and male trait. A second group of models, good-genes or indicator models, suggest that because male traits

are costly to produce and maintain, the expression of these traits is a honest indicator of male quality, and thus females benefit from mating with highly ornamented individuals. Genetic models have been developed that support the plausibility of both Fisherian and good-genes processes. There is also mixed empirical evidence for both processes (see review in Andersson 1994). No study so far has measured the covariance between repertoire size and preference in a single species, which is a primary prediction of both Fisherian and good-genes models (Bakker 1993). However, Searcy (1992) showed in a multispecies comparison that there is a correlation between song repertoire size and the strength of female preference for repertoires. A unique prediction of good-genes models is that there should be a positive correlation between viability and degree of ornamentation. Such an effect has been shown in the great reed warbler *Acrocephalus arundinaceus* (Hasselquist et al. 1996), where offspring viability is predicted by the father's song repertoire. The same authors showed that cuckolded males had smaller repertoires than their cuckolders, and this suggests that the benefit behind EP females' behaviour is increased viability of the offspring.

My goals in this thesis have been to explore several functional alternatives so as to understand the maintenance of song repertoires in the willow warbler (*Phylloscopus trochilus*). The choice for this species was largely arbitrary and prompted by the chance of being able to work on a population which has been colour ringed by Dr. Jim Cobb since 1988. It is know from previous studies that song rate plays an important role in mate choice in this species (Radesäter et al. 1987). However, no thorough research has addressed the extent to which song repertoire is relevant in mate choice and male-male competition. The objectives of the research were:

- to analyse in detail the song complexity of the species, aiming at the identification of those parameters of song complexity which were more likely to be sexually selected;
- to examine to what extent females used these characteristics in mate choice, both at the time of pairing and in the context of extra-pair mating;
- to determine the correlation between male song characteristics and reproductive success, taking into account extra-pair paternity as a component of success;
- to test whether the amount of quality of paternal care could be predicted from the song characteristics;
- to study how the variation in relevant song characteristics influence the aggressive response of territorial males.

Mire vuestra merced - respondió Sancho - que aquellos que allí se parecen no son gigantes, sino molinos de viento, y lo que en ellos parecen brazos son las aspas, que, volteadas del viento, hacen andar la piedra del molino.

Bien parece - respondió don Quijote - que no estás cursado en esto de las aventuras: ellos son gigantes; y si tienes miedo, quítate de ahí, y ponte en oración en el espacio que yo voy a entrar con ellos en fiera y desigual batalla.

M. de CERVANTES

Chapter 1

SONG ORGANISATION AND SINGING PATTERNS

Abstract

A descriptive analysis of the song repertoire of the willow warbler is presented. The song repertoire of this species is hierarchically organised, in that an individual starts all its songs with the same element, and then choices of particular elements lead to predictable sequences. These elements are specific to the individual although some are shared with other individuals. Although the whole element repertoire is difficult to measure (new elements keep on being found even after analysing 100 songs) the majority of elements sung in 90% of songs are already present in the first 15 songs recorded. Principal component analysis performed using a total of 14 variables concerned with song repertoire and complexity yielded 5 components which describe uncorrelated aspects of the song characteristics. The different sequences that compose a bird's repertoire are sung with continuous variety, and a random delivery seems to explain the final output. Reliable estimations of the scores for the five principal components of a given individual can be based on the analysis of only 15 consecutive songs.

Introduction

A thorough description of the behaviour is the first step of any ethological endeavour (Slater 1978). This chapter aims at describing how the song of the willow warbler is organised, and what characterises individual differences. This seems a good way of identifying the possible targets of sexual selection in the song. Bird song is a complex phenomenon, and many levels of analysis are possible. It is known that individuals vary in the rate at which they sing, and in how much they sing; this has been linked to female choice in this species (Radesäter et al. 1987). Also, it is likely that sound frequency variables are important, as we know that in many species minimum frequency is a good predictor of body size (e.g. Schubert 1976). Any form of female choice based on body size could make use of this acoustic predictor. However, I have decided to concentrate on a different level of analysis, that of song complexity.

A common feature of many passerine songs is that they are composed of a repertoire of species-specific elements or song types (Catchpole & Slater 1995). Complex song repertoires have been shown to play in important role in male-male competition and female attraction, and current function suggests that they have evolved through sexual selection in most species (Searcy & Andersson 1986).

Songbirds can be divided in two big groups depending on the way they sing their song repertoires (Hartshorne 1973). The first group is that of the species where each male has a repertoire of several fixed sequences of elements, each one of them is repeated many times before switching to another one. These species are said to sing with eventual variety. A common example of these is the chaffinch *Fringilla coelebs* (Slater 1983). On the other hand, other species sing with immediate variety, that is, each song produced tends to be different from the previous one. In many of the species in this second group song sequences are not fixed, and the constituting elements of the song are produced in a flexible way (i.e sedge warbler *Acrocephalus arundinaceus*; Catchpole 1976).

The first study of the song repertoire of the willow warbler (Schubert 1967) described the presence of several songtypes (*Grundtypen*) in the repertoire of each individual, although these songtypes were not always homogeneous. Schubert did not attempt to measure this variation, and classified songtypes by eye. Other authors have also skipped the study of this variation in their descriptions of the song (e.g. Thielcke & Linsenmair 1963), and looked at either temporal parameters (Helb 1973), or the versatility of the song (i.e. the average ratio of different elements to total elements in each song: Järvi et al. 1980).

It has recently been suggested than a mere count of songtypes or elements might not be the most adequate measure of song complexity (Podos et al. 1992; Nowicki & Podos 1993). This argument has been proposed in the case of the song sparrow *Melospiza melodia*, where there is a considerable degree of variation within each of the songtypes comprising a repertoire. These authors suggested that the degree of between- and within-songtype variability could be more important variables than the number of songtypes. This is so because birds with more songtypes have songtypes which are more similar between them (Podos et al. 1992). However, this hypothesis has not been tested yet in that species. Although the song of the willow warbler does not have the same organisation as that of the song sparrow, it presents a similar problem in that songtypes are not fixed structures, and there are differences in the similarity of songs within songtypes and between songtypes in a repertoire. I have therefore followed the approach of Podos et al. (1992) to some extent.

In this chapter I will describe the organisation of the song of the willow warbler and its pattern of singing. I will also consider several variables of song complexity which may be important in the context of sexual selection, and the way in which these may be best measured. A principal assumption of this analysis is the equal status of the different elements that compose the song. As will be shown latter, songs are composed of different elements, which can be classified by their shape, frequency spectra and length. No attempt was made to analyse whether the physical features of these units were of any relevance, and it is assumed that there are no differences in the cost of singing the different elements, but rather that the important differences lie in the number and organisation of elements. This might be a simplification, given that a recent study has found differences in the relative potency of rapid trills and other song elements for evoking sexual responses in female canaries (Vallet et al. 1998), but it seems most likely that the target of sexual selection in the song of the willow warbler is song complexity per se.

Material and Methods

Subjects and study site

Recordings of birds were obtained in the field in Kippo Wood, (Fife, Scotland), in the springs of 1994-1997. I selected bouts of at least 15 consecutive songs recorded in the period between five days after arrival and the time of mating for a total of 107 different birds. Previous analyses showed that during this interval of time song output is consistent and stable. It also seem likely that 15 songs would be a reliable sample of a male's repertoire. However, in order to confirm this, a sample of 100 songs taken for 10 males was compared to the standard 15-song sample for these males. A further test of the reliability of the 15-song sample was obtained by estimating the repeatabilities of all measurements obtained in a set of 12 males for which more than one sample per bird and year had been analysed.

Song analysis

Visual inspection of the sonagrams generated with a Kay DSP Sonagraph model 5500 or the SIGNAL software was used to prepare a catalogue of different elements. Such a catalogue was built for each of the males. Elements were labelled by letter combinations unique to each bird: no attempt was made to classify elements between different birds. From each song the following direct measurements were taken using the cursors on the Sonagraph screen: song length, post-song interval length and element sequence. A custom-

built program in C (written by Benjamín Pérez Casado) was used to obtain further data from these files: element counts, versatility measures, matrices of element transitions, song similarity coefficients and several other descriptive statistics from the songs. The direct song variables measured were:

- Element repertoire: total number of elements present in the sample of 15 songs.
- Total elements per song: average total number of elements per song.
- Different elements per song: average number of different elements per song.
- Song versatility: average of the ratio of different elements to total number of elements per song.
- Song length: average song length.
- Element rate: average ratio of total number of elements to song length.
- Interval length: average inter-song interval length.
- Consecutive change: average distance between consecutive songs as measured by the Levenshtein distance (see below "Dendrogram construction and analysis" for an explanation of the Levenshtein distance).
- Variants: total number of songs composed of any different sequence of elements, without taking into account consecutive repetitions of the same element.

Transitions between elements

Due to the high number of elements and the quite fixed way in which these are sequenced within the song, the transition matrix contains a very high number of noughts, making it difficult to assess the significance of the sequences (Slater 1973). However, I decided to carry out an analysis of significance on the transition matrixes for practical and descriptive reasons. I followed the method described by van Hooff (1982) which calculates adjusted residuals from every pair-wise sequence. Adjusted residuals are a measure of the degree to which an observed transition frequency deviates from the expected one (van Hooff 1982). A significance value can be drawn for these values, which indicates whether an element facilitates or drives the transition to another one.

Another way in which the fixation of transitions at different points of the song sequence was measured involved the calculation of the uncertainty of a transition from one element to the rest. If an element always precedes another particular element, the sequence is said to have no uncertainty, the transition is certain. If on the contrary, an element can be followed by a large number of elements, the uncertainty increases. The formula of Shannon-Weaver (Steinberg 1977) was used to calculate uncertainties.

Dendrogram construction and analysis

I used an approach similar to that used by Podos et al. (1992) for the song sparrow. These authors used a numerical taxonomy technique (cluster analysis) to describe repertoires in a way that reflects the complexity of the song more fully than a mere count of songtypes. They found that two important parameters showed large degrees of variation between individuals, namely the internal variation within a songtype, and the degree of separation between different songtypes. My general approach was based on their rationale, but I departed in many ways from their methodology, due to the differences between the repertoires of the two species.

The element sequence of all songs was simplified to the basic sequence of elements without consecutive repetitions. Pair-wise distances were calculated between each pair of songs. I chose to use Levenshtein distances rather than the Jaccard similarity coefficient used by Podos et al. (1992). This was done in order to be able to compare birds of different repertoires on the same standard scale. The Levenshtein distance (Kruskal 1983) allows this, because it is the measure of the minimum number of deletions and insertions necessary to go from one sequence to another. I used a program written by Daniel Margoliash for this purpose, in which deletions and insertions were assigned a cost of 1, and substitutions a cost of 2 (see Table 1.1 for an example of how these different coefficients compare). As distance coefficients are needed for the cluster analysis, Levenshtein distances between each pair of songs were divided by 25 (this is an arbitrary number greater than any distance currently present in the set of data).

Comparison	Levenshtein's distance	Levenshtein's distance / 25	1-Jaccard's similarity
ABCDEF – ABCDE	1	0.04	0.167
ABCDEF – ABCDEZ	2	0.08	0.286
ABCDEF – APQRST	10	0.40	0.909
ABC – APQ	4	0.16	0.800

Table 1.1. Four imaginary sequences have been chosen to show the merits of the two distance measurements considered in the text. In this case, Jaccard's similarity has been inverted, so that it can be read as a distance. Note in the third and fourth comparisons how Jaccard's similarity fails to correct for the different sequence length and yields similar figures in the two comparisons, whereas Levenshtein's distance takes into account the overall length of the sequences.

UPGMA cluster analyses (NEIGHBOR application from PHYLIP; Felsenstein 1993) were performed on each of the pair-wise song comparison matrixes for each of the birds. All songs in the sample, rather than only the different variants (cf. Podos et al. 1992) were used in the analysis. The reason for this is that measures of repertoire and cluster homogeneity are taken into account more accurately by considering all songs, as individual willow warblers vary in the degree to which variants are repeated.

However, cluster analysis does not provide an objective way of deciding how many significant clusters exist in the dendrogram. Podos et al. (1992) used a moat index that gives a cut-off point in the dendrogram to reach a final figure for the number of clusters. However, in spite of the appeal of the moat index technique, its use was discarded when it was found that the presence of one very different songs in the sample of a bird would make the moat index peak at fewer clusters, regardless of the clear existence of clusters beyond that point. Willow warblers do not have songtypes like those typical of eventual variety songsters, and thus a cluster analysis of their repertoire does not present the obvious gaps between categories found in Podos et al.'s study. Instead, an arbitrary cut point at a linkage value of 0.8 was used in all the birds (see Fig. 1.1). This value was decided upon after checking that it would yield a reasonable number of clusters. The measurements taken from the cluster analyses are:

- Cuts in tree: This is the total number of branches that are intercepted by the line set at the value of 0.8 linkage.
- Clusters: Number of branches cut by the 0.8 linkage line that contain more than one song.
- Songs per cluster: Average number of songs per branch.
- Repertoire homogeneity: This is 1 minus the maximum linkage value found in the dendrogram. The higher this value, the less diverse the repertoire is.
- Cluster homogeneity: This is 1 minus the weighted average of each of the within-cluster linkage values. These are calculated by averaging the linkage values between adjacent songs in each cluster.

As mentioned earlier, there is no standard way of analysing a dendrogram. The protocol followed in this case responded to the necessity of obtaining a good description of the topology of the tree for those variables of likely biological significance. Although different approaches could be taken, it was felt that the one adopted here responded efficiently to the complexity of the song repertoire. The use of the same measurements for all birds allowed comparisons between individuals to be made.

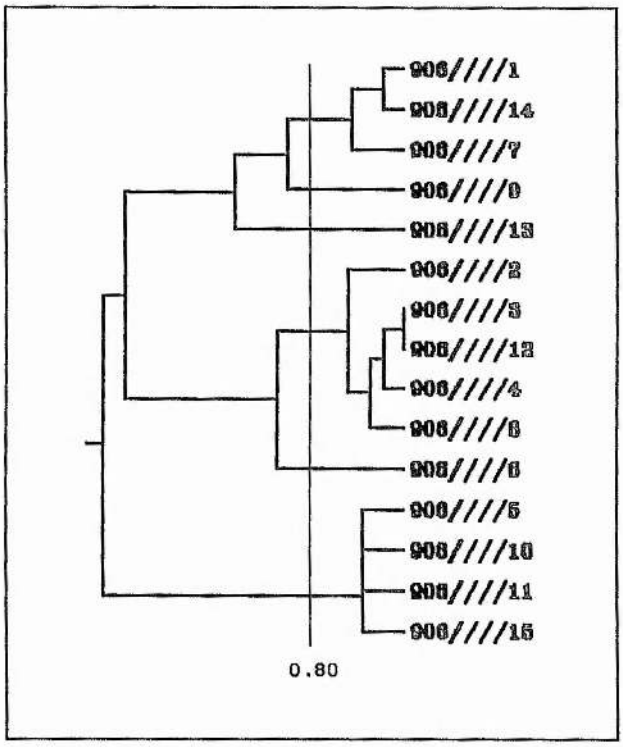


Figure 1.1. Dendrogram obtained using UPGMA cluster analysis on the matrix of similarities of a bird's repertoire. The vertical line at the linkage of 0.8 was used to obtained a comparable number of clusters for all individuals. In this case, the line yields a total of 3 clusters of songs, and 3 individual songs. Homogeneity of clusters was measured by comparing the distance between adjacent songs within each cluster.

Principal component analysis

Although the use of many parameters seems appropriate when describing a complex phenomenon like birdsong, many of these variables are inevitably intercorrelated and measure similar attributes. A solution to this problem is to extract from the set of variables a few factors that account for the greater part of the variance in the data (Frey & Pimentel 1978). Principal component analysis was performed on the 14 variables measured in the data set. The interpretation of the components took account of the sign and loadings of the different variables. In all cases, a biologically meaningful label could be assigned to the components.

The minimum sample size

One of the aims of this chapter is to provide a firm basis for the quick assessment of variables of repertoire complexity in the field, which will be related to functional aspects. I

found that a sample size of 15 songs is easily obtainable for a large number of birds, but larger samples may not always be easy to record. Spearman rank correlations were performed between all song and repertoire variables measured in the sample of 100 songs and in a limited one of 15 songs, to assess how well these samples would predict the real repertoire.

A complementary analysis of this issue was performed by calculating the repeatabilities of all song variables and components considered. Repeatability is also known as the "intraclass correlation coefficient" and is a measure of the within-individual variance in relation to that between-individuals (Sokal & Rohlf 1995). A high repeatability value indicates that the character being measured has a high stereotypy in each individual, and that differences between individuals are high enough for selection to act on that character (Boake 1989).

Patterns of singing

In order to describe the sequences of songs, I looked at the average distance between songs at different lags from each other. I used the Levenshtein distance for this purpose, and measured the difference between songs at lags 1, 2, 3... 16 (where lag 1 means consecutive songs; lag 2, two songs away...). This was done for four birds for which a large number of songs could be recorded and analysed. In order to provide a measure of the departure from what random singing would yield, bootstrapping analysis was conducted on the whole similarity matrix between all songs in the sample. By using the distance matrix of all sampled songs (i.e. many of them repeated several times), rather than a matrix of all different variants, it is assumed that not the whole repertoire is accessible at the moment of singing, and that some songs are more likely to be produced than others.

Results

Basic song description

Hall-Craggs in BWP (1992) describes the song of the willow warbler as a "*lyrical drooping melody of gentle, pure notes in silvery, rippling phrase*"; this is too good a description not to be quoted. Each song lasts about 3 seconds and is composed of individual elements, each of them normally repeated several times before switching to a new one (Fig. 1.2). The range of frequency is normally from 2 to 7 kHz, and the pattern of delivery typically drops in frequency: the first elements are very high pitched, and the following ones descend quickly to the song minimum (Fig. 1.2).

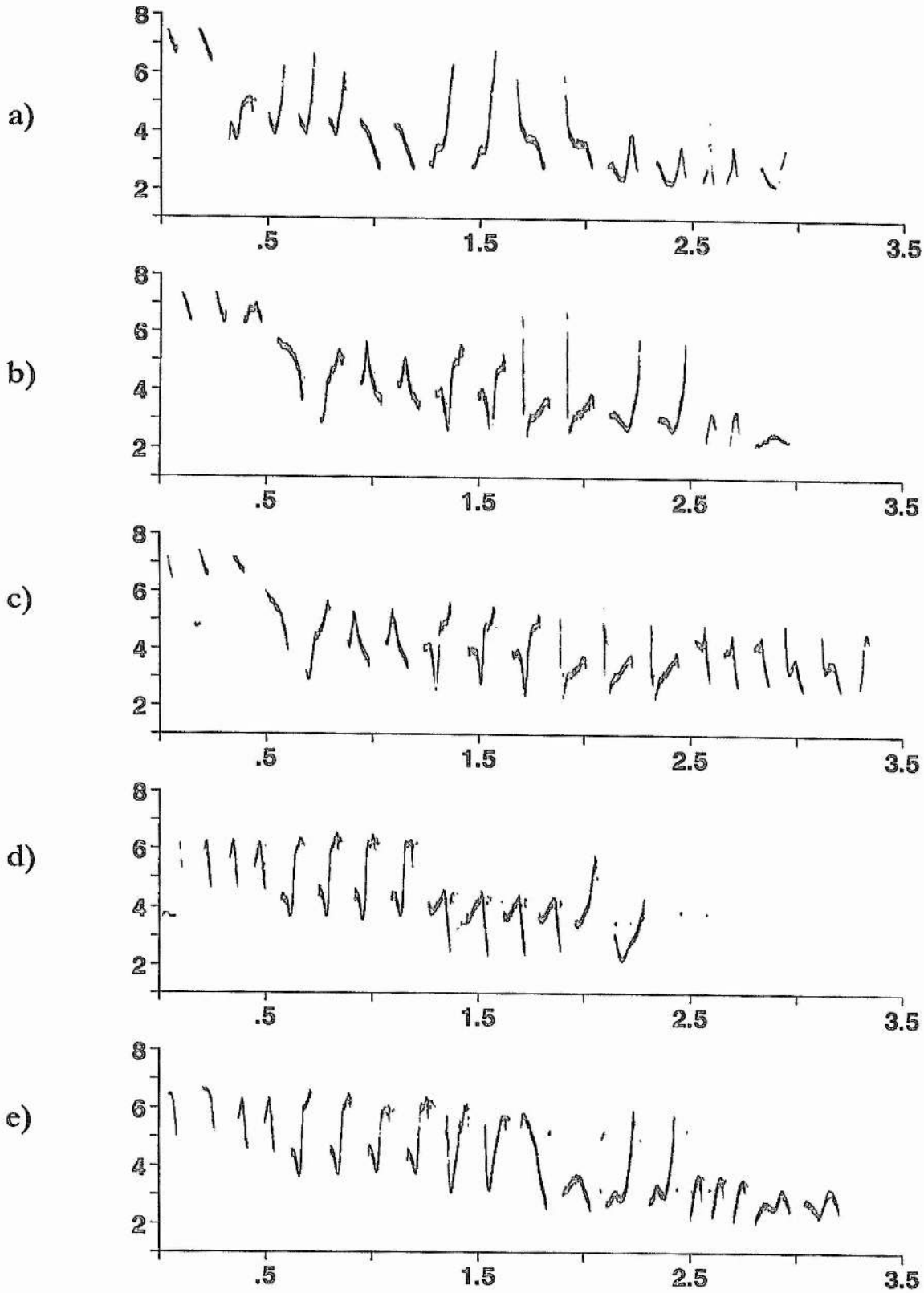


Figure 1.2. Three songs (a-c) from bird #97, and two (d-e) from bird #13. Note how some elements can be found in different songs of the same individual (e.g. the first three different elements in songs d and e are the same ones).

Songs are rarely repeated consecutively, and new elements are produced as we listen to new songs. It is possible to plot new elements as they appear in the sample of a particular bird (Fig. 1.3) in order to assess its element repertoire (Wildenthal 1965). Some of the birds clearly reach an asymptote before 100 songs have been examined (see Fig. 1.3) and, thus, a final element repertoire size can be given. However, for others the cumulative plot keeps on increasing and it is difficult to estimate where the ceiling may be. If the lines were asymptotical, it would be possible to estimate the limit, but new elements do not seem to appear randomly, but in groups, and this makes it impossible to apply the methods suggested in the literature (Catchpole & Slater 1995).

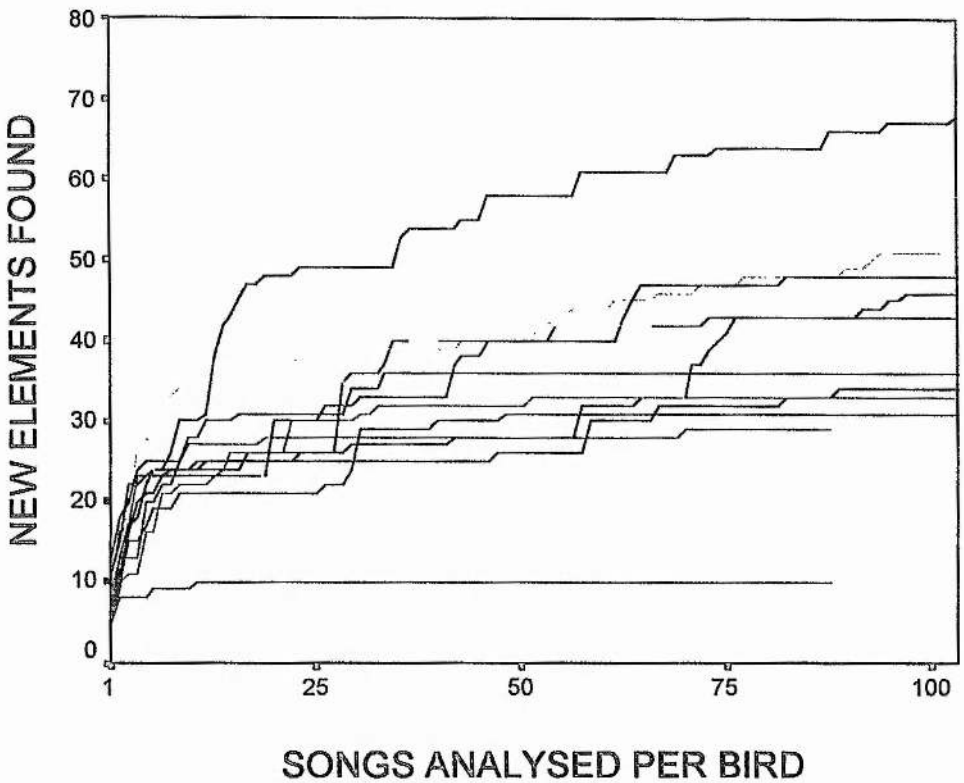


Fig. 1.3 Cumulative plot showing how the number of new elements produced increases with number of songs analysed. The graph shows the data available for all 12 males considered. Note how the sample of 100 songs fails to give a final ceiling for most of the birds, although some show a clear asymptote.

It is interesting to look at the coefficient of variation (CV) between individuals in the different variables measured in the song, as this will indicate which variables might be of importance in the context of sexual selection. Three different measures of song repertoire have the highest CVs among all the variables: number of variants, element repertoire and consecutive change (Table 1.2).

Variables	Mean	SD	CV	Min	Max
Element repertoire	29.69	9.38	31.61	9.00	67.00
Total elements per song	17.02	1.76	10.35	11.60	21.93
Different elements per song	8.30	1.26	15.18	6.00	12.27
Song versatility	49.03	6.79	13.85	34.15	64.77
Song length	2.95	0.28	9.46	2.17	3.64
Interval length	1.76	0.18	10.13	1.18	2.23
Consecutive change	10.35	3.14	30.35	1.14	18.36
Variants	51.37	18.61	36.22	6.80	89.80
Element rate	5.77	0.34	5.81	4.95	6.74

Table 1.1 Summary data for the nine direct variables drawn from the samples of 15 songs from 107 different birds. The descriptive statistics have been obtained from the mean values for each of the birds.

Sequencing of elements

As can be seen from Fig 1.2, some elements are repeated in several songs of the same individual. Sequences between elements in a bird's repertoire are not random, but highly predictable, and trees of decisions (*sensu* Dawkins & Dawkins 1973) can be built based on the subset of sequences with the highest significance values for the adjusted residuals (Hooff 1982). If we plot the significance values obtained for the adjusted residuals of all possible sequences between elements, a definite bimodal distribution is found (see for example the results for bird #2 in Fig. 1.4). A similar pattern was found in all birds that were examined (10 birds, 100 songs each). Transitions between elements are either very unlikely or almost certain. Taking this last subset of elements, and selecting only the very frequent elements, we can construct trees of transition between elements (Fig 1.5).

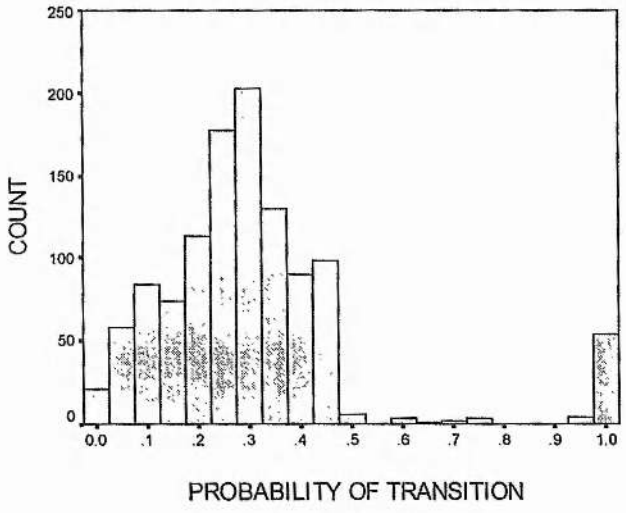


Fig 1.4. Frequency distribution of the probabilities of all possible transitions between the elements of a bird's repertoire (Bird #2 in this case). The probabilities are drawn from the adjusted residuals of each transition (van Hooff 1982).

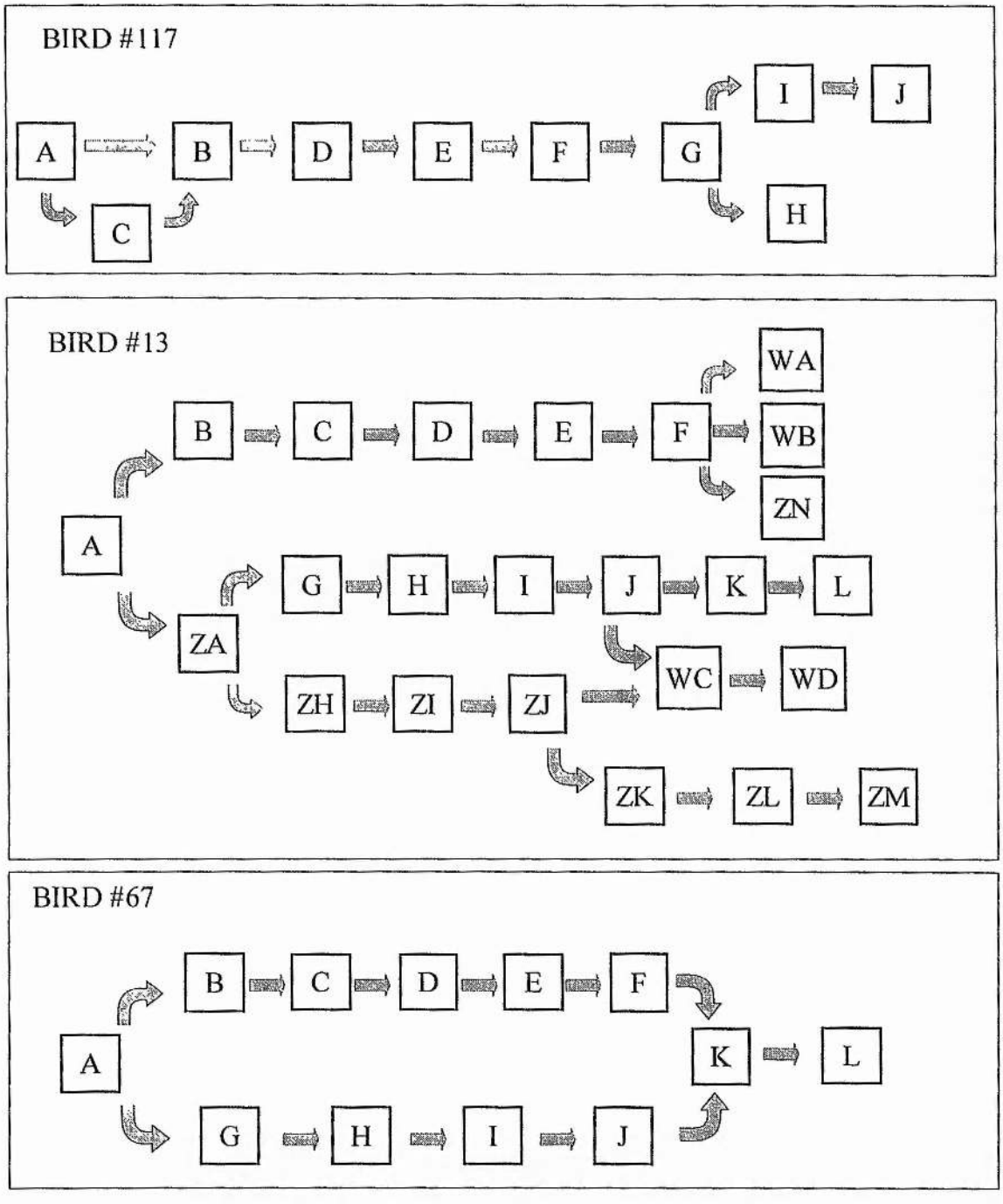


Fig 1.5. Schematic representation of transitions between the elements in the repertoires of 3 different males. Only transitions with a very high significance of occurrence were used, and from these, only common ones were included for simplicity. Letters do not represent the same elements across individuals. See how all birds start with a same element at the beginning, and from here on a branching pattern emerges. Bird #13 is representative of the common pattern in most birds with larger repertoires. Bird #67 presents an unusual pattern, in which the initial branching converges at the end of the sequences. Bird #117 is unusual in having a very limited repertoire. (Element repetitions have been omitted for simplicity)

Typically, individual elements appear at a certain position within the song string (consecutive repetitions not considered). There are introductory elements which always start every sequence, middle part elements, which follow the first, and elements which invariably close sequences. However, there is variation in how fixed elements appear in a given position, and birds show different degrees of element fixation (Fig. 1.6).

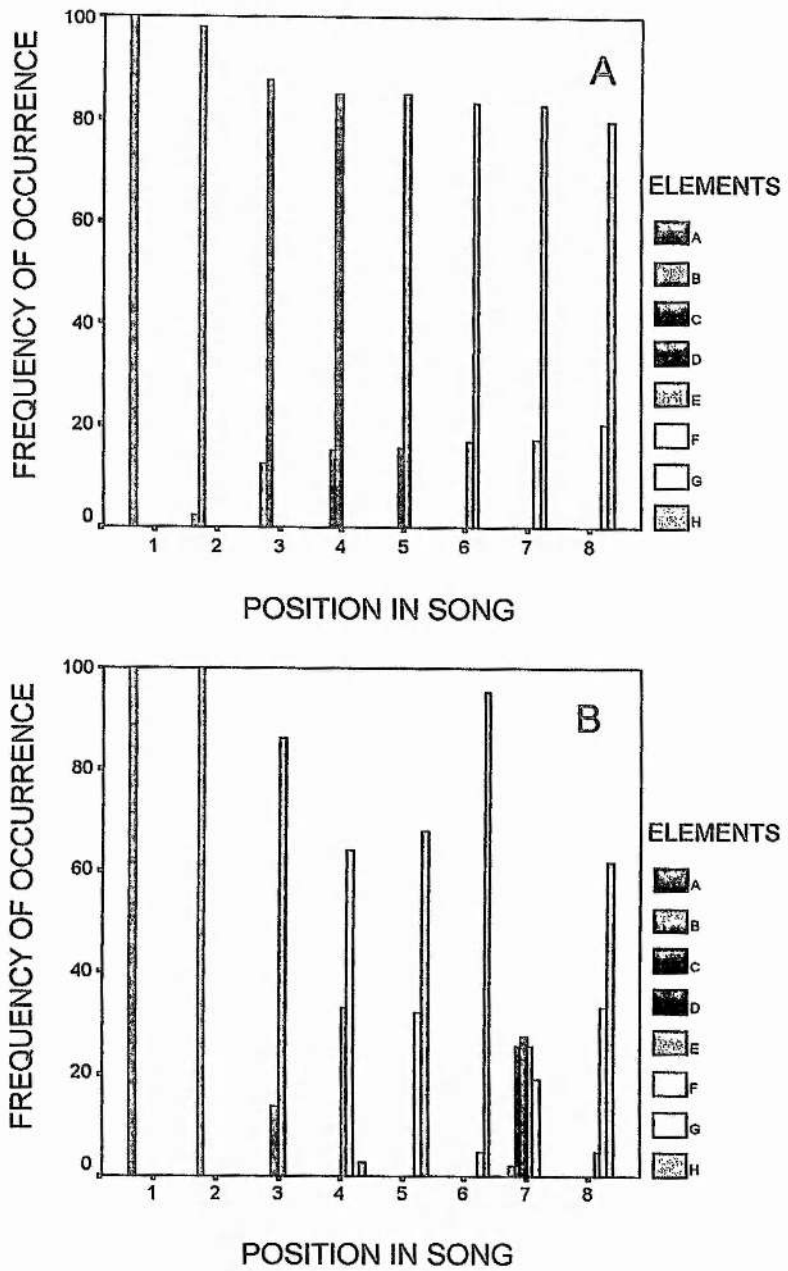


Figure 1.6. Elements are usually produced at a certain position within the song, but there exist differences between birds. These graphs show the frequency of occurrence at different positions of a subset of elements of two birds with extreme values on this respect. Bird 013 (a) has a high fixation level of elements, whereas bird 043 (b) presents a much more lax fixation. Position is measured without considering consecutive repetitions of elements.

As can be seen from the arrangement of the element sequences in Fig. 1.5., some elements lead to almost obligatory transitions, whereas others have several possibilities. A good way of measuring this characteristic is by using the of Shannon-Weaver uncertainty formula (Steinberg 1977). High values in uncertainty indicate that a transition from an element is not certain, i.e. it is difficult to predict which element will follow. Low figures on the contrary indicate that we can predict the next element in the sequence with a high degree of confidence. Given the particular hierarchical nature of the song organisation (Fig 1.5.), it follows that the uncertainty of transition varies with position in the song, as points of decision do not appear randomly along the sequence, but rather in two main areas. To measure this, I took the twenty most frequent elements for each of the birds, and calculated their average uncertainty of transition. Each element was assigned a position in the song based in the modal point of its occurrence, and averages were calculated for all elements pooled. (Fig. 1.7).

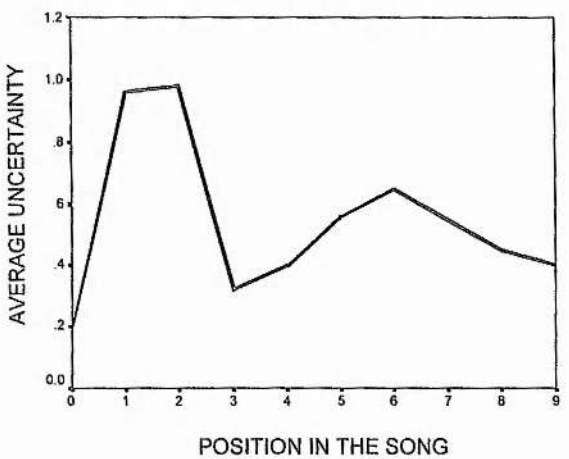


Fig 1.7. Average uncertainty throughout the song sequence, averaged for the 20 most frequent elements of 10 birds. Changes in uncertainty reflect the hierarchical organisation of elements shown in Fig. 1.5. The two peaks of uncertainty correspond to two major areas of points of decision in the structure of transitions in the song. Position 0 refers to the uncertainty before any element is sung.

Uncertainty is very low at the beginning, as all birds start their songs with the same element, and increases greatly in the second position, when several key elements can follow. After this, uncertainty decreases again, as these elements lead to rather fixed sequences. The two peaks of uncertainty correspond to positions in the song sequence which are characterised by the presence of decision points (Fig. 1.5; see also Dawkins & Dawkins 1973).

Use of elements

Not all elements are sung equally frequently. There are roughly three different types of elements according to the frequency with which they appear in a birds' repertoire: always present, alternatively present and rarely sung (Fig. 1.8).

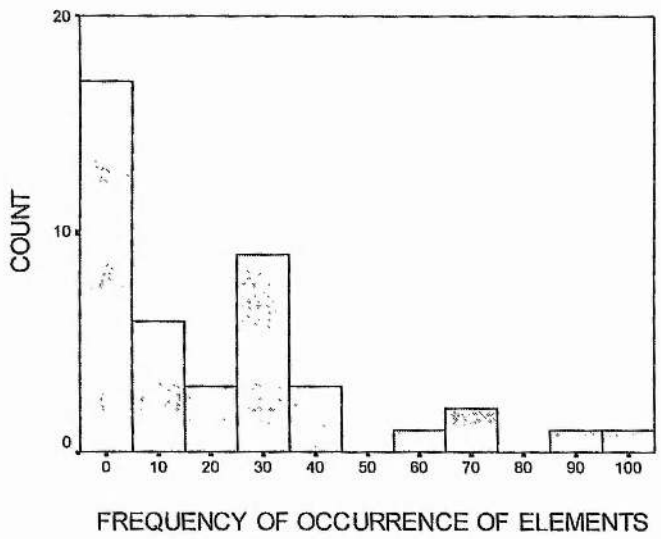


Figure 1.8. Frequency distribution of the occurrence of particular elements in a single bird's repertoire. Note how the distribution can be loosely divided in three sections: elements that are always present, frequent elements, and rare ones. Figures along the x axis correspond to the lower limit of the interval of each bar.

An important characteristic of the song of this species is that, although the element repertoire seems very large (see Fig 1.3), some elements are repeated very often, and these form the core of the repertoire. This group of elements appears again and again throughout the sample, and although new elements can appear, these are rarely sung again (see Fig 1.8). To test this point, I took all elements that occurred in at least 10 songs in the sample of 100 songs studied in 10 males, and checked whether these elements could be found in a restricted sample of 15 consecutive songs. The size of this core of elements ranged between 21 and 30, and on average 95.22% (SD=3.68, n=10) of them were sung in the sample of 15 consecutive songs. This means that, although new elements can always appear, the bird's core repertoire does not change depending on the time it has been recorded.

There is also a very strong relationship between frequency of occurrence and position in the song (Fig. 1.9), with the most frequent elements sung at the beginning of the song. The general pattern is that introductory elements are invariably the same for each bird, and they thus appear in almost every song sung by that individual. These are also shared to a greater extent across individuals. These elements are normally very high pitched and are repeated

several times, showing a *crescendo* pattern. They may perhaps be warming up elements. After these follow those that appear in the mid frequency range, which are the main bases of the different sequences, as we shall see later. Finally, we have those that appear in even smaller numbers at the end of the songs. Given the hierarchical way in which song is organised, it follows that those elements at the end of the sequence are be the less frequent ones.

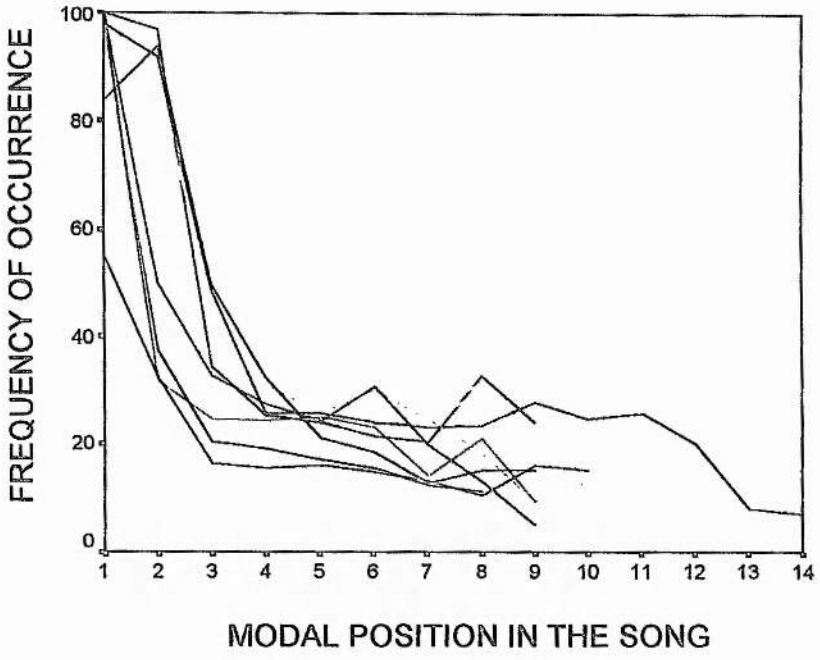


Figure 1.9. Commonness of occurrence changes with position in the song in the 10 birds examined. In the graph, each different element has been allocated a position within the song according to its statistical mode. Each line represents a bird, and each data point is the mean frequency of occurrence for all elements within their typical position. The pattern is clear: the commoner elements are sung at the very beginning of the song, those less common in the middle section and rarer ones towards the end.

Higher order repertoire structure

A visual analysis of the dendrograms produced by the UPGMA cluster analysis revealed that similar songs were indeed clustered together. A moat index such as the one used by Podos et al. (1992) could not be used, as the branching of the clusters occurred across the whole width of the tree in many birds, leaving no obvious gaps like those found in song sparrow repertoires (compare birds 4 and 62 in Figure 1.10). This led to the use of the arbitrary cut-off point explained in the Methods section.

Podos et al. (1992) suggested a possible constraint in the organisation of the song of the song sparrow, by which an increase in the number of songtypes was achieved at the expense that there was also an increase in the similarity of the different songtypes. In the case of the willow warbler, however, the number of clusters in the dendrogram

("songtypes") was not positively correlated with the repertoire homogeneity ($r=-0.170$, $p<0.080$, $n=107$). In other words, larger repertoires were not less diverse. At any rate, the analogy with the song sparrow work should not be taken too far, because there are major differences in song organisation between the two species.

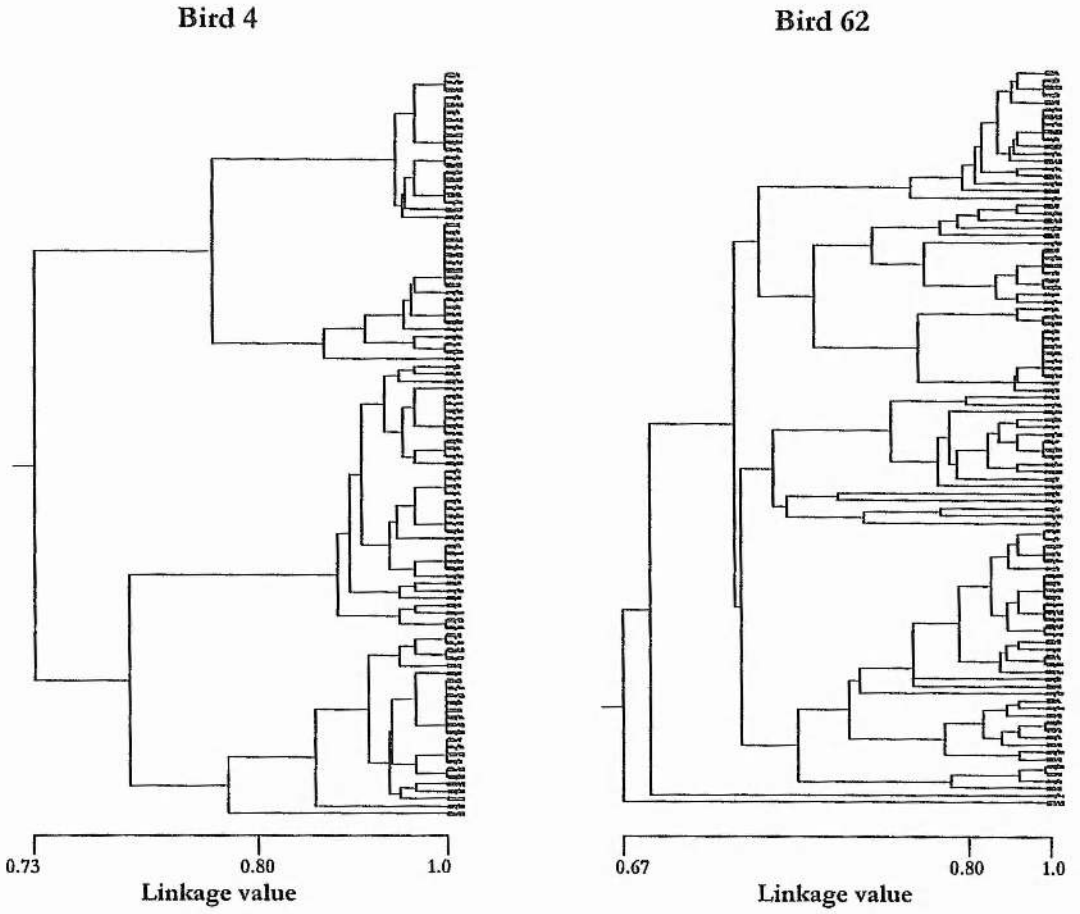


Figure 1.10. Trees produced by UPGMA cluster analysis on the song similarity matrixes of two birds with different repertoire structures. Bird 4 has a highly structured song, with highly defined clusters, whereas bird 62 has a poorly structured repertoire, with many songs that do not fall into clear clusters. The x axis is the linkage value, 0.80 being the cut-off point used to obtain an unbiased estimate of number of clusters.

Multivariate analysis of song variables

The principal component analysis performed on the 14 direct variables (Table 1.3) yielded five principal components that account for a 82.93% of the total variance (Table 1.2).

The nature of these components can be unveiled by looking at the weightings of each variable (Table 1.3); components have been named according to these variables (Table 1.2). The variables that contribute to PC1 are all related to repertoire size: element repertoire,

cuts in tree and number of variants contribute positively to scores for this component, and repertoire homogeneity and songs per clusters contribute negatively. It seems reasonable to consider PC1 as measuring element repertoire. PC2 reflects song length, as total elements per song and song length are related to this component. Another component of singing performance is PC4, to which element rate is positively related and song length negatively related. Interestingly, only song versatility contributes heavily to PC3. The last component (PC5) is particularly interesting as it measures the degree of fixation of clusters, independently of other variables: birds with high scores in PC5 have many highly homogeneous clusters of songs, and this suggests that PC5 is a measure of repertoire fixation.

	Eigenvalue	% of variance	Cumulative %
PC1 Song repertoire	5.46	39.00	39.00
PC2 Song length	2.20	15.70	54.70
PC3 Song versatility	1.72	12.31	67.01
PC4 Element rate	1.20	8.57	75.58
PC5 Repertoire fixation	1.03	7.35	82.93

Table 1.2. Eigenvalues and variance explained by the five first principal components extracted from the set of 14 original variables. Each component has been labelled after analysis of the contribution of the weightings of the different variables (Table 1.3).

Variables	PC1	PC2	PC3	PC4	PC5
Element repertoire	0.880	-0.129	-0.068	0.088	0.026
Cuts in tree	0.844	-0.396	-0.249	-0.014	-0.044
Songs per cluster	-0.844	0.396	0.249	0.015	0.045
Repertoire homogeneity	-0.804	-0.264	-0.257	-0.033	-0.126
Consecutive change	0.744	0.290	0.147	0.303	0.182
Variants	0.668	-0.419	-0.029	-0.183	-0.039
Different elements per song	0.639	0.569	0.435	-0.124	0.004
Total elements per song	0.421	0.742	-0.487	-0.042	-0.112
Song length	0.323	0.607	-0.500	-0.476	0.217
Song versatility	0.412	0.077	0.867	-0.103	0.100
Element rate	0.223	0.341	-0.060	0.693	-0.562
Clusters	0.320	-0.210	-0.249	0.510	0.643
Cluster homogeneity	-0.742	0.327	0.025	0.246	0.417
Interval length	-0.232	0.152	-0.116	0.106	-0.016

Table 1.3. Weightings of the original set of variables in the five principal components extracted from them. Weightings of relative importance within each component have been shaded. When the same variable correlates in the same direction with several components, only the highest contribution has been highlighted.

Patterns of singing

Willow warblers sing with continuous variety, and the same song is rarely repeated twice in a row. Given the large number of song variants, and the difficulty of classifying these into songtypes, sequences of songs have been studied by measuring the Levenshtein distance between consecutive songs (see Methods for a definition). This was done for long bouts of songs, and the distance between songs was measured at lags of 1, 2...15 songs. This is effectively an autocorrelation analysis, only that similarities are measured instead of recording recurrence of identical sequences. In order to be able to compare these values to a random pattern, the matrix of distances between all recorded songs was bootstrapped 100 times for each of the sample sizes. Only data for 4 birds are presented, as the pattern is similar in all cases (Fig. 1.11).

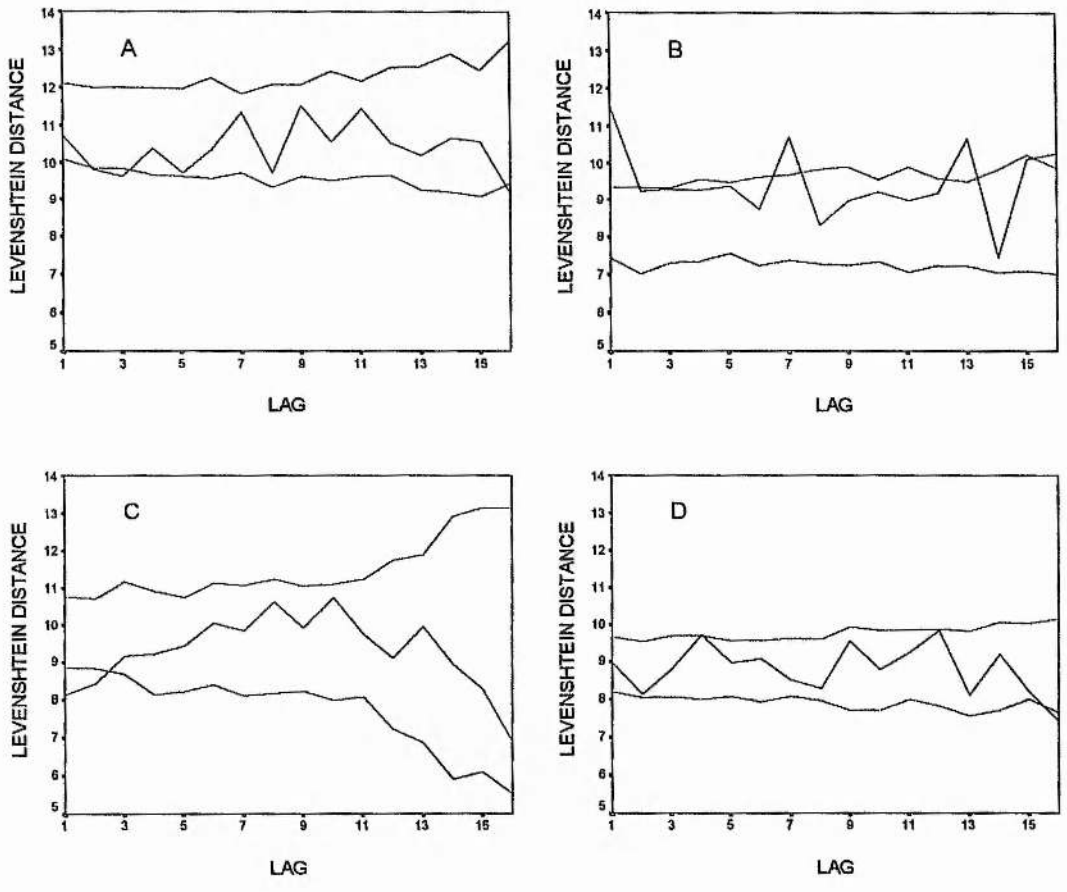


Figure 1.11. Lag-similarity graphs for the song sequences of four individuals. Lines in blue are the 95% outer limits of the expected values. In other words, any red line which crosses the blue boundary in any direction is showing a significant difference with a random pattern. Sample sizes decrease with increasing lags, hence the widening pattern of the confidence intervals.

The lag-similarity graphs for all birds coincide in showing that the general pattern is random. In other words, given the fact that the average distance between songs can be

as low as 0, or as high at 15 in most birds, the average distance at most lags is not too different from a random choice of songs. Within this general randomness, there are still some other patterns, but these seem to be individual-specific. Bird B, for instance, seems to maximise the distance between consecutive songs, as consecutive songs (lag=1) are less similar than expected at random. Bird C on the contrary tends to repeat similar songs in succession more often than expected. Very little evidence of cycles can be seen, although this analysis does not preclude the possibility that specific transitions between individual song variants exist. What is clear is that, even if such transitions do exist, the final output of the bird is one of immediate variety, which in terms of mechanisms could probably be described by a random choice of song sequences.

Repeatability of song variables and principal components

In order to test the reliability of the sample of 15 songs, I calculated the repeatabilities (Lessells & Boag 1987; Sokal & Rohlf 1995) of all song and repertoire variables, as well as principal components. To this effect, an ANOVA with individuals as groups was performed on a sample of 12 birds, six of which had been sampled twice, and four of them three times. The results show very high repeatabilities for all direct song variables except interval length. High repeatabilities are also found in the repertoire variables, with the exception of within cluster homogeneity. The same results are reflected in the principal component scores, with high repeatabilities for all components except repertoire fixation.

Variables	R
Element repertoire	0.721
Total elements per song	0.736
Different elements per song	0.797
Song versatility	0.879
Song length	0.517
Interval length	0.227
Consecutive change	0.818
Variants	0.474
Element rate	0.761
Cuts in tree	0.767
Songs per cluster	0.768
Repertoire homogeneity	0.717
Clusters	0.827
Cluster homogeneity	0.219

Principal components	R
Song repertoire (PC1)	0.815
Song length (PC2)	0.516
Song versatility (PC3)	0.854
Element rate (PC4)	0.699
Repertoire fixation (PC5)	0.215

Table 1.4. Same year repeatabilities (R) for all song and repertoire variables and for principal components scores.

A different approach to this problem can be taken by looking at the concordance of the results obtained from the sample of 15 and 100 songs obtained for 10 birds, and calculating the correlation between the two sets for each variable (Martin & Bateson 1993). The Spearman correlation coefficients are shown in Table 1.5, and they coincide with the repeatability values in that the two variables with the lowest reliabilities are interval length and cluster homogeneity.

Variable	r_s
Element repertoire	0.763
Cuts in tree	0.512
Songs per cluster	0.723
Repertoire homogeneity	0.813
Consecutive change	0.515
Variants	0.789
Different elements per song	0.842
Total elements per song	0.891
Song length	0.723
Song versatility	0.903
Element rate	0.782
Clusters	0.611
Cluster homogeneity	0.394
Interval length	0.418

Table 1.5. Reliability of the samples of 15 songs compared with those of 100, as shown by Spearman's correlation coefficients for the different variables.

Discussion

The description of the willow warbler song repertoire outlines above can be summarised as follows:

a) the elements that compose a song are not randomly put together, but sung with rather few possible transitions between elements;
b) the organisation of elements corresponds to a hierarchically branching pattern, where elements either connect predictably to a second one, or constitute a decision point where two or more available transitions are possible;
c) this branching pattern pervades the whole length of the sequence of elements, and suggests that the concept of songtype is not appropriate in this species;

d) generally speaking, consecutive songs are not more similar from each other that if they were delivered at random;
e) there is a high degree of variability between individuals in several characteristics related to the size of the song repertoire.

The concept of songtype is a very dear one to birdsong researchers. It is usually applied to all species that show more than one species-specific song pattern. However, its use *sensu stricto* refers to those species that produce different and repeatable song patterns, which can be easily classified into different categories notwithstanding small variations. Species of this kind are the chaffinch (Slater 1981) or the song sparrow (Podos et al. 1992). In the case of those species that build their repertoire of more or less freely recombined elements, like the sedge warbler, the use of the concept of songtype is not appropriate, as new recombinations of elements are continuously produced (Catchpole 1976). Elements, instead, are the building blocks of the song repertoire (Buchanan & Catchpole 1997).

In the case of the willow warbler, we are halfway between those two extremes. In the diagram of element sequences (Fig 1.5), we can see that the branching progresses from the second element, and points of decision appear at different points later in the sequence. We can decide on a cut-off point for practical reasons, but this is an arbitrary choice. If we look at the repertoires of birds #117 and #67 (see Fig. 1.5) we might consider that they have two song types. By bird 13 standards, those songtypes would seem minor variations of one single one. The element, rather than the songtype, is likely to be the unit of production and memory in this species. Hence the avoidance of using the term songtype; cluster is a more neutral term and does not imply underlying mechanisms of production and perception.

However, some individuals do present something very close to a songtype repertoire like those typical of other species. These are those birds with many clusters and high cluster homogeneity, (i.e. birds with high scores in PC5: repertoire fixation). It can be hypothesised that a development towards this pattern is brought about by age. This kind of organisation with more deterministic sequences may be no less efficient than a highly variable one, specially if the degree of consecutive change is kept at a maximum. Besides, a highly organised repertoire is likely to be much less costly in terms of neural capacity. This is worth studying, as it fits with theories on the development of cost-reducing traits that can help the maintenance or production of secondary sexual characters (Møller 1996). In this context, a large repertoire could be less costly to produce and store if it had a high degree of fixation.

Sequences between songs were not analysed with respect to the identity of each song, and only the similarity between them was considered. This method would not detect possible cycles between specific sequences, such as those found for other species (e.g. Todt 1970; Lemon & Chatfield 1971; Slater 1983). Individual patterns were found, and for instance, from the lag-similarity plot for individual B (Fig. 1.11), an inhibition on similar songs seems to be set after each song is sung, whereas in individual C a facilitation is more likely to occur. Inhibition of a particular songtype after production has been described in other species (i.e. the varied thrush *Zoothera naevia* ; Whitney 1981), but in this case the effect seems to be nearly negligible, and individually specific. The general pattern found was that consecutive songs were not more similar or different than if they were delivered at random. This singing pattern is interesting from a functional point of view, and some speculation about its design is worthwhile. We expect organisms to display their sexually selected characters in a way that facilitates a quick assessment by the receiver (Sullivan 1994). If it is adaptive for a male bird to be perceived as possessing a large repertoire, it would be advantageous to sing this repertoire in a way that maximises the *perceived* size of the repertoire. This might be specially true in species where repertoires are of limited size, like the willow warbler. A singing pattern which is random in terms of similarity between consecutive songs (ABBDCABDDC...) is likely to tax more heavily the memory capacities of a receiver which is estimating its size than, for instance, one in which the bird would cycle through the repertoire in a fixed way (ABCDABCD...). The *perceived* repertoire size might effectively be enlarged by exploiting a possible receiver psychological constraint, in terms of memory capacity in this case (Guilford & Dawkins 1991). Note that this only applies to species with smallish repertoires; in the case of large repertoires (e.g. the nightingale *Luscinia megarrhynchos*; Todt & Hultsch, 1996), highly deterministic sequences could ensure that never the same song can be remember by a receiver. Another consequence of this singing pattern is that habituation would be minimised, as there are no consecutive repetitions of songs or sequences of songs. An interesting experiment could be done by testing whether receivers (both males and females) are sensitive to this aspect of song organisation.

The first principal component (PC1) was interpreted as measuring song repertoire. All variables that measure different aspects of repertoire size, like number of variants, element repertoire or degree of consecutive change contribute to it, and it is important to note that this component has been clearly isolated from the rest by these characteristics. We can now use it to rank individuals along the underlying common factor of song repertoire, which is not correlated with the other variance components, and relate it to functional aspects. It is

interesting to note that those variables that had the highest contribution to this component also had the highest variation between individuals (Table 1.1). This variation, together with its high same-year repeatability score should make PC1 an obvious target for sexual selection (Boake 1989).

Repeatability is a measure of the proportion of phenotypic variance of a trait which is due to consistent differences between individuals, and sets an upper limit to heritability (Falconer & Mackay 1996). However, a measure of repeatability is highly dependent on the time scale in which this is taken. This time scale should be relevant to the organism. In the case of the willow warbler, between-year repeatabilities are a more reliable estimate (Chapter 4), given that the degree of survival for adults is as high as 50% (Chapter 5).

The constraint suggested by Podos et al. (1992) in the song of the song sparrow was not found: birds with larger repertoires did not have less variability between songs. Apart from the different methodology used to measure their attributes, the two species have very different song organisation systems, and thus constraints on their repertoire sizes are likely to be different. However, a similar rationale could be made with the fifth principal component mentioned above, which can be interpreted as a measure of repertoire fixation. The aim of principal component analysis is to identify uncorrelated factors, and thus the relationship between these two components can only be determined by looking at their changes between years in the same individuals (Chapter 3).

One of the main aims of this chapter was to provide a reliable way of measuring repertoire characteristics in the field in order to look at its functional aspects. The high correlations between the 15-sample and the 100-sample sizes for most variables suggest that 15 consecutive songs recorded 5 days after arrival and before pairing provide a good measure of song repertoire characteristics. This point was also confirmed by the high repeatability scores. Some variables like interval length had much lower repeatabilities, and they fall into what Gerhardt (1992) has labelled "dynamic properties" of acoustic signals. This "dynamism" does not preclude high repeatabilities at the same time, and directional sexual selection preference has been shown for some dynamic attributes of calls in treefrogs (Gerhardt 1992). The fact that several measures of 15-songs have low repeatability does not mean that that attribute does not have a high repeatability if, for instance, environmental and social conditions were taken into account (Boake 1989), or if female assessment is based on longer times than the three minutes it takes to sing 15 songs. For instance, it is known that song rate is an important aspect of mate choice in this species (Radesäter et al. 1987), and between-song interval length contributes to this variable.

However, song-rate is probably assessed in long bouts and thus my small sample size has failed to find a relevance of it.

At a more mechanistic and rather out of fashion level, it is worth considering the results of Fig. 1.7. This shows that the graph of predictability against time is a discontinuous one for the song of this species, as also observed by Dawkins & Dawkins (1973) for the drinking behaviour of chicks. The advantage of looking at the song of this species is that we can define clear elements of production (i.e. elements) which must have a highly specific neural path. Dawkins & Dawkins (1974) argue that behavioural descriptions of this kind can pave the way for neuroethological investigations. It would be fascinating to see how the neural structure of the song nuclei reflects the song organisation described here.

Chapter 2

SINGING ACTIVITY AND MATE GUARDING

Abstract

In most bird species, males use a whole set of behaviours which reduce their chances of being cuckolded. The 'fertility announcement' hypothesis proposes that song might also act as a paternity guard (Møller 1991). According to this idea, paired males would announce their female's fertile status by singing. This has been interpreted as an honest signalling, evolutionarily stable strategy. I tested some predictions from this hypothesis in the willow warbler. Contrary to the predictions, males sang very little when females were fertile. Intrusions by other males were not less common when males sang at higher rates. Mate guarding and singing are best interpreted as two conflicting behaviours, the former being directed to the fertile female and the later to the attraction of a second one. A revision of recent studies suggests that in most passerine species, males do not sing during the fertile period of their females, except at dawn. The opposite results found by Møller (1991) could be due to his use of population-wide estimates of singing behaviour and egg-laying. Asynchronies in individual breeding cycles could be responsible for the apparent matching between peaks of singing activity and fertility in a population.

Introduction

Male birds spend a considerable proportion of their time in close proximity of the female during her fertile period (Birkhead & Møller 1992). This behaviour has been described for many species, and its function is best explained as guarding the female from fertilisation attempts by other males (Birkhead & Møller 1992). Mate guarding normally involves behaviours like following the female, staying for long periods of time in her vicinity, frequent copulations, territorial defence, cloaca pecking, etc. (Birkhead & Møller 1992). Møller (1988) suggested that song also had a mate guarding function, and that the evolution of this behaviour could be a case of honest signalling (Møller 1991). According

to this hypothesis, only good quality males would sing at high rates during their female's fertile period. Cuckolder males would maximise their reproductive success by not intruding into territories in which the owner male sang intensively, as females paired to these males would be unlikely to accept EPCs. Two predictions from this hypothesis are that males should sing at higher rates when their female's fertility is at its highest, and that intrusion rates by other males should be inversely related to the song rate of the owner of the territory.

The function of a behaviour is best demonstrated by experimental manipulation. For instance, that song serves to defend territories has been shown by muting territorial males and looking at their success in keeping a territory (McDonald 1989). However, another possibility is to look at the context in which singing occurs and correlate song occurrence with different situations of biological relevance (Catchpole 1982). Although this approach cannot conclusively prove a link between function and behaviour, it is still a basic way of understanding behaviour. Even if correlation does not imply causation, lack of a correlation between two variables makes a causal link very unlikely.

I have recorded singing rates, mating and mate guarding behaviour in a population of willow warblers in order to look at how song is used in the context of mate guarding. Singing and guarding behaviours were analysed with reference to the precise stage of the female's fertile period.

Methods

The study area was a mixed woodland of spruce, birch and willows in the Kippo estate (Fife, Scotland), where willow warblers breed at a high density. All birds were colour ringed. Territorial males were observed daily to map their territories, and precise settlement and pairing dates were taken for all males. Between the arrival and nestling stages, focal watches of 10 min were conducted on territorial males. Observations took place in the first four hours after dawn, as willow warblers do not sing in the dawn chorus. To avoid a sampling bias towards more easily observable males, a random 'target' list was drawn every day, trying to obtain a similar number of observations for each of the breeding stages. A total of 21 males were observed at some of the breeding stages, but not all them could be observed for any given stage. On each day, a targeted male was looked for, and the focal observation would start five minutes after he was located. If a bird could not be found after 10 min of searching, the next male in the list was looked for.

During these focal observations, the following behavioural categories were recorded: number of songs per minute, occurrence of alarm or contact calls, copulation attempts or copulations, feedings by the male and intrusion by other males. Distance from the female was difficult to assess due to the density of vegetation, and I only recorded whether the female was within 5 m of the male or not. If the focal bird disappeared from sight, the focal watch was stopped until the bird could be found again. If this did not happen in four minutes, the observation was stopped altogether.

Nest building is done only by the female, and nests were found by following females on the pairing day and on the three following days, when nest building activity is at its highest. Nests were checked every two days in order to obtain the date of laying of the first egg. Only males for which the date of the first egg laid by their females was known were considered in this study. Focal observations became difficult during the fertile period, as very often neither the member of the pair could be found. This precluded a day-to-day analysis of song and behaviour throughout the season. Instead, I divided the breeding season into the following stages:

1) four days after territory settlement and before pairing;
2) pairing day (the first day when a male was seen in proximity to the female that eventually bred with him);
3) fertile period (from 6 days before the first egg was laid until that day; this was based on what is known for other passerines (Birkhead & Møller 1992) and from a previous study of copulation behaviour in the willow warbler (Arvidsson 1992). This period does not overlap with the pairing day;
4) laying period (the 6 days after the first egg was laid, and before incubation);
5) incubation (only the female incubates);
6) feeding nestlings (both parents feed the young).

Form the observed behaviours, only song rate had a normal distribution. Statistical analyses were thus parametric for song rate, and nonparametric for the rest of variables. Behavioural categories like copulations, intrusions, calling and feedings occurred at a very low rate, so data were considered in a categorical fashion as occurring or not in a given focal observation. Individual data for these behaviours were then transformed as the percentage of focal observations for a given individual and period during which a given category of behaviour was registered. When individual averages or frequencies for a certain period are considered, only males with a minimum of 3 focal observations in that period were included in the sample (average 3.76, SD=1.34).

Results

Singing rates differed greatly between the different stages (ANOVA; $F=20.86$, df=5,69; $p<0.001$; Fig. 2.1). The highest rates were found before pairing, and the lowest on the day of pairing and during the fertile and nestling stages. The day of pairing was characterised by a very pronounced drop in the singing rate, as it has been described before for this species (Radesäter et al. 1987). Song rate during the fertile period was very low, and very different from the high rates of unpaired birds. During laying and incubation, song rates increased again, but did not reach the levels of unpaired birds.

Post-hoc comparisons revealed that song rates of unpaired birds were significantly different to the rest of the stages (Tukey's tests, all $p<0.02$). Song rates during the pairing and fertile stages were not different statistically (Tukey's test, $p>0.05$). However, song rate during the laying period was significantly greater than during the fertile one (Tukey's test, $p<0.01$).

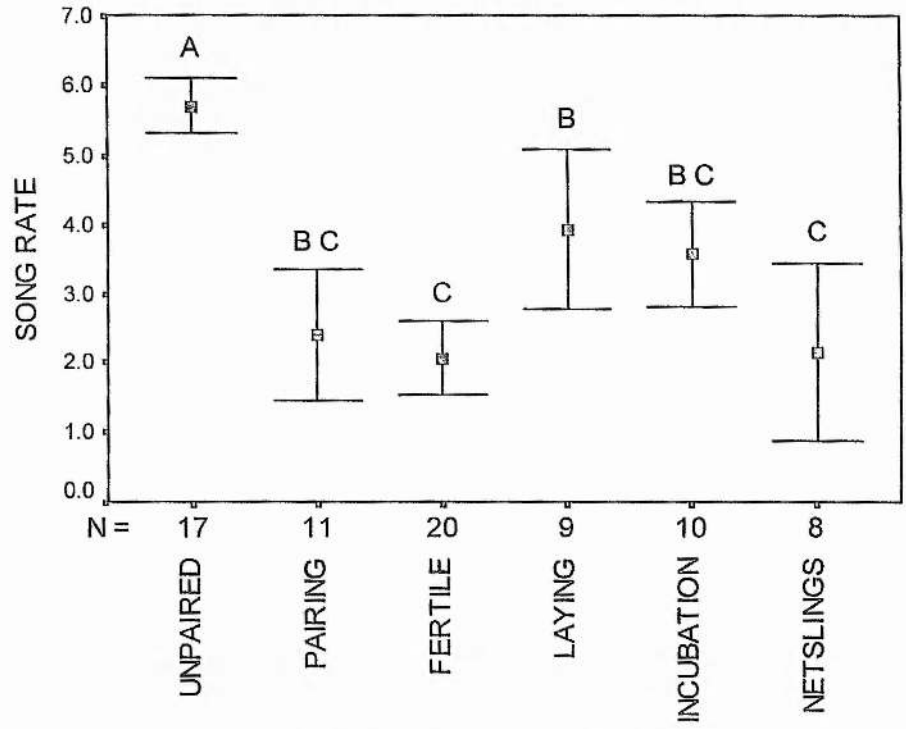


Figure 2.1. The graph shows changes in song rate (mean ± 2SE) throughout several stages of the breeding season. Fertile period was defined as the week before the first egg was laid. Within a given period, data from each individual were averaged. Sample size is given along the x axis. Letters over bar represent homogeneous subsets of groups for illustration (Tukey's HSD test).

Males were observed in proximity to the females in only three periods: day of pairing, fertile period and egg laying. The proportion of time that males spent within 5 m of the female was not the same in these three periods (Kruskal-Wallis; $\chi^2=10.07$, df=2, $p<0.007$; Fig. 2.2). I tested differences between groups using the multiple comparisons test for the

Kruskal-Wallis analysis (Siegel & Castellan 1988): there were no differences between pairing day and fertile period ($p>0.05$), but males spent less time with females once laying started ($p<0.05$).

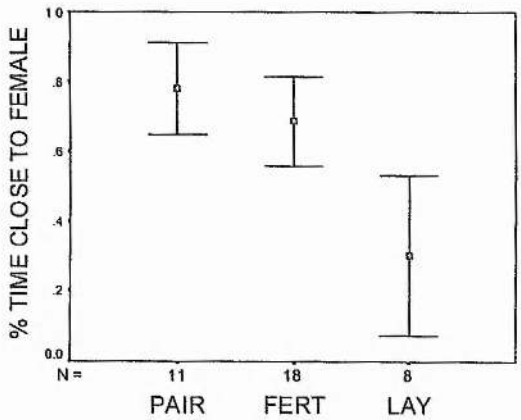


Figure 2.2. Males spent more time within 5 m of the female during the day of pairing and the female fertile period than during egg laying. Graph shows means ± 2SE.

Intrusions by other males into the focal male's territory were different between the different periods (Kruskal-Wallis: $\chi^2=12.71$, df=5, $p<0.02$; Fig 2.3). Although none of the multiple comparisons between periods was significantly different (all $p>0.05$), the trend is one of higher rates of intrusions during the fertile period and on the day of pairing.

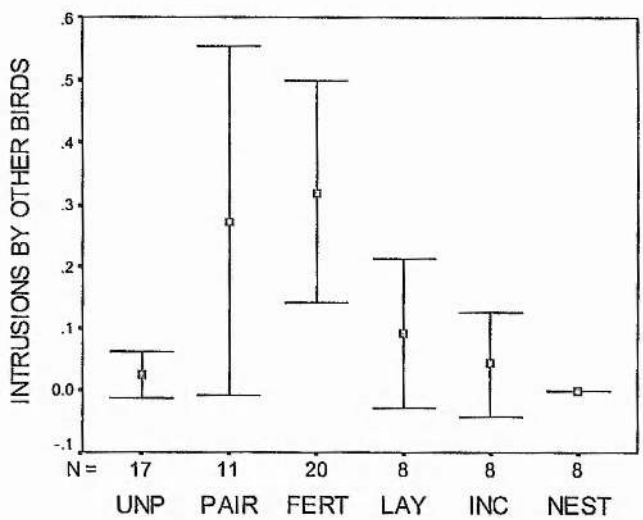


Figure 2.3. Frequency of focals with intrusions by other males throughout the breeding season. Abbreviations are: UNP: unpaired; PAIR: pairing date; FERT; fertile period; LAY: laying period; INC: incubation; and NEST: nestlings. Graph shows means ± 2SE

Table 2.1 shows averages of the percentage of focals per period in which several behaviours were observed. These were not very frequent, so statistical tests were only performed for the overall differences between the four periods. Among the general patterns that can be observed it is interesting to note a high level of alarm calling by the

male during the laying period, and also frequent contact calling by the female during the fertile period and on the pairing day. Copulations peaked during the fertile period, and weaned thereafter.

Behaviour	Pairing day	Fertile period	Laying	Incubation	χ^2	p
Copulation attempts	0.18	0.11	0.04	0.03	0.728	0.113
Copulations	0.55	0.65	0.31	0.00	7.346	0.064
Feedings	0.09	0.15	0.00	0.00	5.564	0.135
Alarm calls	0.09	0.01	0.20	0.00	11.28	0.010
Contact calls	0.64	0.32	0.04	0.00	12.52	0.006
n	11	23	27	25		

Table 2.1. Frequency of occurrence of five different categories of behaviour during several periods related to mate guarding. Data represent averages of the individual percentages of focal observations in which the behaviour was registered. Statistical tests were Kruskal-Wallis, all with 3 degrees of freedom.

Considering the periods when the male could be close to the female, that is between the pairing date and the egg laying periods inclusively, there was a negative correlation between song rate and percentage of time spent within 5 m of the female (r_s=-0.474, n=21, p<0.030; Fig 2.4).

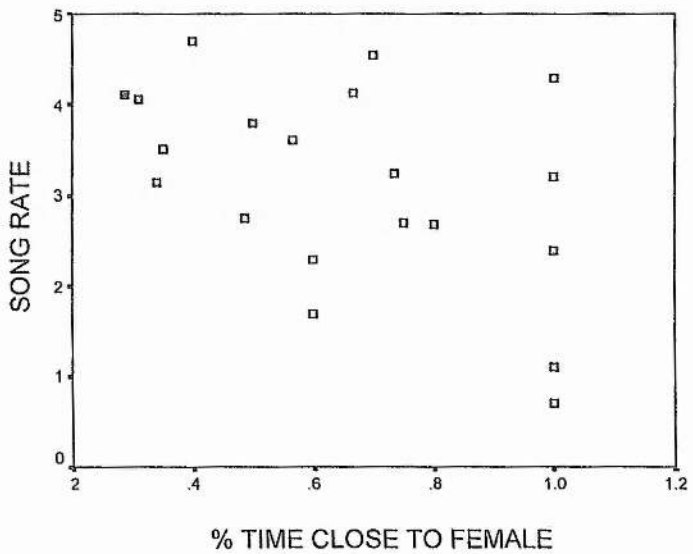


Figure 2.4. There is a negative correlation between song rate and percentage of time spent within 5 m of the female (r_s=-0.474, n=21, p<0.030).

One of the predictions of the fertility announcement hypothesis (Møller, 1991) is that the rate of intrusion by other males should be negatively correlated with the song rate of the territory owner. Contrary to this prediction, the correlation between the average song rate during the fertile period and the occurrence of intrusions was a positive one, albeit non significant (r_s=0.446, n=12, p=0.146). Overall, song rates were not lower when intrusions occurred than when they did not (Wilcoxon matched pairs test, Z=-0.178, n=12, p=0.859), but the percentage of time males remained near the female was greater in those focal observations in which intrusions were recorded (Wilcoxon matched pairs test, Z=-2.193, n=12, p=0.028).

Another prediction of the fertility announcement hypothesis is that high quality males should announce the fertility status of their females at a higher level than low quality ones. In willow warblers, song rate is a sexually selected character (Radesäter et al. 1987), so we would expect to find a positive correlation between song rates before pairing and during the fertile period. Quite on the contrary, there was an almost significant negative correlation between song rates before pairing and during the fertile period (r_s=-0.560, n=12, p<0.058). Considering individual males, the frequency of focals in the fertile period in which intrusions were recorded was negatively correlated with the song rate of the male before pairing (r_s=-0.612, p<0.034, n=12). That is to say, high quality males (as judged by earlier song rate) were less likely to suffer intrusions, and this was not related to higher singing rates during the fertile period. This pattern could be due to high quality males using the alternative tactic of staying closer to the female for a higher percentage of the time, but there was no correlation between song rate before pairing and percentage of time within 5 m of the female during the fertile period (r_s=0.185, n=12, p<0.565).

Discussion

Three main results from this study run counter the predictions of the fertility announcement hypothesis (Møller 1991):

1) Paired males greatly reduced their singing rates when their female was fertile, and singing rates rise again during laying and incubation.
2) Males with high song rates before pairing did not sing comparatively more when their female was fertile, the trend rather being in the opposite direction.
3) Intrusions by other males were not less likely to happen when the territory owner was singing at high rates. On the contrary, the correlation between singing rate and intrusion was a positive one.

The observed pattern suggests that mate guarding and singing at high rates are two conflicting behaviours subject to a trade-off. In this population, some males attract a second female, and there is a high level of EPP (20% of nestlings, 45% of clutches). The second peak of singing after the female's fertile period is likely to function to attract second females or extra-pair ones. A similar picture has been described for great reed warblers *Acrocephalus arundinaceus* (Hasselquist & Bensch 1991). In that species, the chances of attracting a second female decrease with date, and males forego mate guarding and start singing long songs to attract a second female only early in the season, when their chances of getting this second female are high.

It is difficult to interpret the negative correlation found between time spent near the female and singing rate. This could be due to temporal variations in fertility, males singing when the potential success of an EPC would be relatively low. Alternatively, males might not be able to mate guard at times because they cannot find their mate, and thus singing might be a way of contacting her. This second possibility was suggested by Rodrigues (1996) to account for similar results found in the chiffchaff *Phylloscopus collybita*. He interpreted the lower song rates when the members of pair where together as suggesting that males used singing as a way of contacting the female. Yet other possibility is that males may only attempt to attract a second mate (or other females for EPCs) when not with their own female. These hypotheses are not mutually exclusive, and they all fit with the conflicting nature of the two behaviours.

It is usually assumed that only males benefit from mate guarding. However, the behaviour of extra-pair males may constitute a considerable harassment and mate guarding can be advantageous to females (e.g. Kempenaers et al. 1995). The fact that males with high singing rates before pairing suffered fewer intrusions suggests that females paired to these males suffered less harassment by EP males. How could this difference come about? It could be that these males are better at mate guarding, and that they use more effective strategies. Alternatively, EP males may direct fewer EPCs attempts towards females paired to high quality males. This second explanation would fit Møller's hypothesis (1991) if these high quality males sang at higher rates during the fertile period. But the results show the opposite: high quality males sang comparatively less during the fertile period. This rather suggests that high quality males had fewer intrusions as a result of their behaviour during mate guarding or before pairing. Although I did not find that high quality males spent a greater percentage of time at close distance from the female, other form of vigilance might have been used. Alternatively, EP males might use song rates before pairing as a cue to a male's quality and the likelihood of his female accepting EPCs.

The fact that males used more alarm calls during the laying period suggests that females may benefit from defence from predators at that time. There is a small percentage of cases of intraspecific brood parasitism in the population. Male vigilance may also protect females from being spied on by prospective parasitic females.

The patterns of mate guarding observed in this study are very similar to those described by Arvidsson (1992) for a Swedish population of the same species. As in that study, most successful copulations happened during the fertile period. The average distance between the two birds increased significantly once egg laying started, suggesting that keeping close to the female is a mate guarding strategy. This pattern was also found in the Swedish study. Intrusions by other males peaked both on the day of pairing and during the fertile period. In contrast, Arvidsson (1992) observed very few intrusions on the day of pairing. Differences in breeding density between the two populations might account for these differences in extra-pair male behaviour. It is interesting to note that no extra-pair copulations were observed in the present study, although there is a high level of EPP in this population.

A number of studies have attempted to test the fertility announcement hypothesis since its first formulation by Møller (1991). As can be seen from table 2.2., only two of these studies have found data apparently supporting the hypothesis. This is in contrast to the original literature survey conducted by Møller and his own data on the yellowhammer *Emberiza citrinella* (Møller 1988). In the literature survey (Møller 1991) 71.4% of the 49 species surveyed presented a peak in singing activity during the female's fertile period. How can these two sets of data be reconciled?

Many of the studies used by Møller (1991) in his survey seem to rely on population-wide data for singing and egg-laying data. As no bird population breeds in perfect synchrony, it follows that any generalisation from population data to individual strategies will blur time patterns. This is specially true if something as precisely timed as the fertile period is considered. The most likely explanation for the apparent fit of the peak of singing with that of fertility found by Møller (1991) is a result of this lack of synchrony.

One of the main functions of song is mate attraction (Catchpole & Slater 1995). This is known from experimental studies (i.e. Eriksson & Wallin 1986) and from correlational observations which show that males of some species stop singing when they obtain a mate (e.g. Catchpole 1982). Given that not all males get paired at the same time, the peak of singing of any songbird population will tend to coincide with the fertile periods of the

females of the first males to get paired. In contrast with the studies used by Møller (1991) to support his hypothesis, most of the recent studies considered in Table 2.2 aimed at directly testing whether males sang during the fertile period, and used known individual fertility data.

The only study which has found higher singing rates when females are fertile since 1991 is during the dawn chorus of the willow tit (Welling et al. 1995). Singing in the dawn chorus can only be considered a paternity guard if males are able to physically guard the female at that time, and females can also obtain EPCs during that time. However, these high singing rates are observed before dawn and while the female is roosting inside the nest cavity (Welling et al. 1995; Otter et al. 1997). Similar results found in the blackbird *Turdus merula* (Cuthill & Macdonald 1990) and in the American robin *Turdus migratorius* (Slagsvold 1996), also suggest that a peak in dawn singing is related to female fertility, but song is not used after the female emerges from roosting. Dawn chorus in these species can be understood as a continuous male assessment situation (Otter et al. 1997), and only as a second-best strategy when direct mate guarding is not possible.

Species	Were individual fertile periods known?	More song when female is fertile?	Reference
Ficedula albicollis	YES	NO	(Pärt 1991)
Luscinia svecica	YES?	NO	(Merilä & Sorjonen 1994)
Fringilla coelebs	YES	NO	(Sheldon 1994)
Parus montanus	YES	YES (at dawn)	(Welling et al. 1995)
Luscinia svecica	Partially	NO	(Krokene et al. 1996)
Emberiza schoeniclus	NO?	NO	(Nemeth 1996)
Phylloscopus collybita	YES	NO	(Rodrigues 1996)
Junco hyemalis	YES	NO	(Titus et al. 1997)
Troglodytes aedon	YES	NO	(Johnson & Kermott 1991)
Oenanthe oenanthe	Partially	NO (song) YES (flight displays)	(Currie et al. 1998)
Phylloscopus trochilus	YES	NO	this study

Table 2.2. Results from the studies which address directly the question proposed by Møller (1991) of whether song increases during the female's fertile period.

There is an overwhelming lack of support for the fertility announcement hypothesis. It is still possible that in very open habitats, where mate guarding is easily done without close contact, male birds might not forego singing after pairing in order to attract more females or EPCs partners (i.e. in the stonechat *Saxicola torquata* (Greig-Smith 1982) or in the wheatear *O. oenanthe* (Currie et al. 1998). The case of the yellowhammer (Møller 1988) is peculiar, as conventional mate guarding seems to be almost absent in the species (Sundberg 1992).

To conclude, as it stands now, there is no convincing evidence that singing is an effective means of mate guarding in any bird species.

Chapter 3

CHANGES IN SONG CHARACTERISTICS WITH AGE

Abstract

Repertoire size, song versatility and element rate increased between the second and third year of age in male willow warblers. No appreciable change in any song characteristic was found beyond that age. A possible trade-off between different characteristics of song was found, in that there was a tendency for increases in song repertoire to be related to decreases in average song length. Arrival date could be predicted from several song characteristics, and this pattern seemed to change with age. This could be explained if there is variation in the cost of different song characteristics with age.

Introduction

Good-genes models of sexual selection propose that because male ornaments reduce viability, their expression must reflect a male's phenotypic or genetic quality (Sutherland & De Jong 1991; Andersson 1994). Amongst other things, it should pay males to display traits that advertise age to prospective mates, to the extent that survival is heritable, so that offspring of long lived males also tend to survive for long (Trivers 1972). Even if longevity is not heritable, females mating with old males may secure direct benefits. For instance, old males may provide better parental care through improved foraging or defence against predators (Alatalo et al. 1986; Marchetti & Price 1989). Females may thus use ornaments as indicators of age, though it is important to note that a positive correlation between age and the expression of sexually selected characters can arise both through differential mortality and by an increase in the size of the ornament with age (e.g. Manning 1989).

In many songbirds, females have preferences for males with a large song repertoire (Searcy 1992), and repertoire size has been found to be positively correlated with viability in the

great reed warbler *Acrocephalus arundinaceus* (Hasselquist et al. 1996). A correlation between age and repertoire size has been shown for several songbirds in the field (Yasukawa et al. 1980; Järvi 1983; Cosens & Sealy 1986; Lambrechts & Dhondt 1986; Hiebert et al. 1989; Eens et al. 1992; Lampe & Espmark 1994; Birkhead et al. 1997). These studies are cross-sectional studies of different age cohorts, so it is also possible that the observed pattern is due to an increased survival of birds with large repertoires. However, individual between-year increases in repertoire size have also been found in several species (Nottebohm & Nottebohm 1978; Espmark & Lampe 1993), which indicates that birds can either learn late in life, or start to sing songtypes or elements memorised earlier (Catchpole & Slater 1995). Evidence suggests that possessing a large repertoire is costly in terms of the brain space involved in storing it (Canady et al. 1984; Devoogd et al. 1993), and there are reasons to believe that increase in size of a particular brain area is achieved to some extent at the expense of others (Jacobs 1996). This makes repertoire size a good candidate for an example of a handicap that can indicate male quality.

The expression of some song characteristics can be affected by developmental constraints, and an increase in one characteristic may be at the expense of decreasing others (Podos et al. 1992). Alternatively, these changes could indicate the expression of a cost-reducing trait which allows an increase in the expression of a sexually selected character (Møller 1996). Bird song is a multivariate phenomenon, and the study of trade-offs between multiple variables may help us to identify the different selective pressures which affect different aspects of song.

Arrival date is an important component of fitness in transaharian migrant passerines (e.g. Møller 1994). Early birds can settle in the most productive territories or those with better nesting sites (Lundberg & Alatalo 1992) and thus have earlier nesting dates, which is a character under directional natural selection (Verhulst & Tinbergen 1991). Arrival date has low repeatability in this species (Chapter 5), and this suggests that it is a condition dependent trait. If a given characteristic of song confers a handicap to males, a negative correlation between that characteristic and arrival date could be expected. A further complication arises when age is taken into account. The costs associated with song are likely to be greater for second-year birds than for older ones. This is so because it is expected that the costs of learning a repertoire for the first time will outweigh those of maintaining and updating one that has already been learned. If this is the case, the expected negative correlation between song repertoire and arrival date should disappear or be reduced with age.

I have looked at changes with age in a set of song characteristics (see Chapter 1) in the willow warbler. The specific question I would like to answer is whether information about age and arrival date, which could be potentially be used in female choice, is conveyed in the song of the male.

Methods

Song characteristics were measured for each male from recordings of 15 consecutive songs, recorded at least 5 days after arrival and before pairing. Five independent principal components were extracted from a larger set of variables measured in the song recording (Chapter 1). These principal components were 1) song repertoire size; 2) song length; 3) song versatility; 4) element rate; and 5) repertoire fixation. Principal component scores are standardised, that is their mean is 0 and the standard deviation is 1.

Once willow warblers fledge, they can only be aged in their first few months, before migrating to Africa. Young birds are characterised at this stage by fresh primaries and a yellowish tinge on their breast and underparts (Svensson 1984; Jenni & Winkler 1994). Notches in their tail feathers can also help to tell juveniles from adults after these have moulted (J.L.S. Cobb, pers. obs.). However, once the birds come back after the first winter, it is impossible to tell them apart from older birds. This is because the willow warbler is one of the two species of passerines in the Paleartic in which both juveniles and adults undergo a complete moult in their winter quarters (Stresemann & Stresemann 1966; Underhill et al. 1992).

Regular ringing of adults and juveniles has been done in Kippo Wood since 1985 by J.L.S. Cobb, as the area is a ringing site within the constant effort ringing scheme (CES) organised by the BTO. This meant that many adult birds were already ringed and their ages known when the study began in 1994. New birds without rings arriving to breed in 1994 were not included in the analysis. After 1994, however, given the fact that most birds come back to the same territories, or very close to where they were in the previous year, I consider all new arriving males as second-years.

Throughout this chapter, I refer to birds in their first breeding season as second-years (2YR). Similarly, third-years (3YR) are birds breeding for the second time. A third category, birds older than third-years (>3YR), was formed to include the remainder of the birds; very few were older than four-years (2 five-years and 1 seven-year).

The data set allowed two different forms of analysis: a cross-sectional one, where age cohorts are compared between each other, and a longitudinal one, where the same birds are compared in consecutive years. A complicating factor of the cross-sectional analysis is that differential mortality may confound the results. For instance, older birds might not have larger repertoires, but birds with larger repertoires could be more likely to survive. In this analysis, for the sake of data independence, individual birds are considered only once. In order to increase the sample size of older ages, if several years were available for one individual, the last year this bird was recorded was taken for the analysis.

The longitudinal analysis involves within-individual comparisons between years. The results of this analysis can help clarify the outcome of the cross-sectional analysis. The largest sample size that could be obtained was for birds recorded both in their 2nd and 3rd years (n=25). Only 5 birds were recorded in their third and fourth years, and these were not included in the analysis due to the small sample size.

Results

Cross-sectional analysis

Analysis of variance showed differences between the age classes in two of the five song principal components, namely repertoire size and versatility (Table 3.1).

	F	p
Song repertoire size	3.818	0.025
Song length	0.761	0.470
Song versatility	5.169	0.007
Element rate	0.372	0.690
Repertoire fixation	1.310	0.275

Table 3.1 Results of the analysis of variance performed for the five song principal components between the three age classes (df=3, 95).

Comparisons between the three age classes revealed that for both song repertoire size (Fig 3.1a) and versatility (Fig 3.1b) the only two groups which differed significantly where 2YR and 3YR (Tukey's tests: repertoire: $p<0.023$; versatility: $p<0.005$); differences between the other groups were not significant ($p>0.05$).

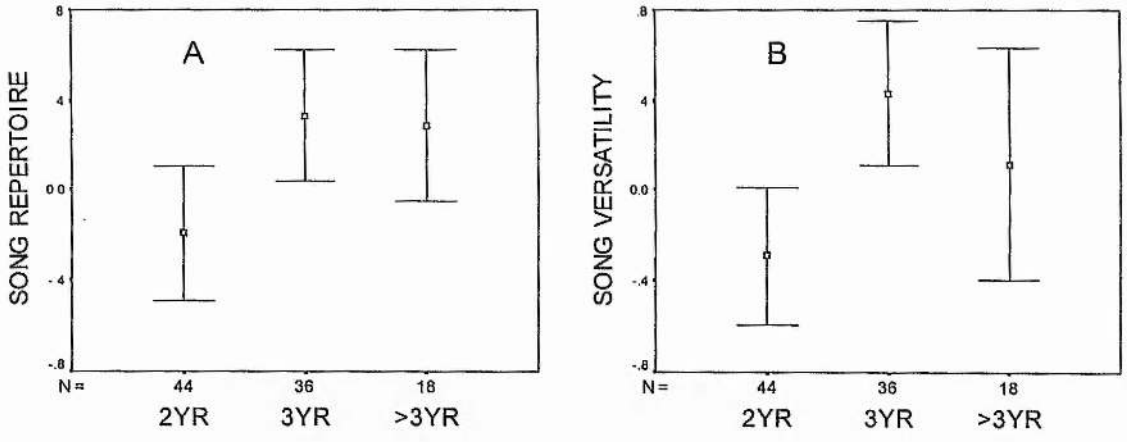


Figure 3.1. Changes in song repertoire size (a) and song versatility (b) between the three age classes. Significant differences exist only between 2YR and 3YR years for both variables.

An analysis of variance showed differences in arrival date between the three age classes ($F=7.604$, $df=2,91$, $p<0.001$). Multiple comparison tests revealed that the groups that differed most were 2YR and >3YR (Tukey's test: $p<0.001$), the rest of the differences between groups not being significant (Tukey's tests; 2YR vs. 3YR: $p=0.070$, 3YR vs. >3YR, $p=0.122$; Fig 3.2).

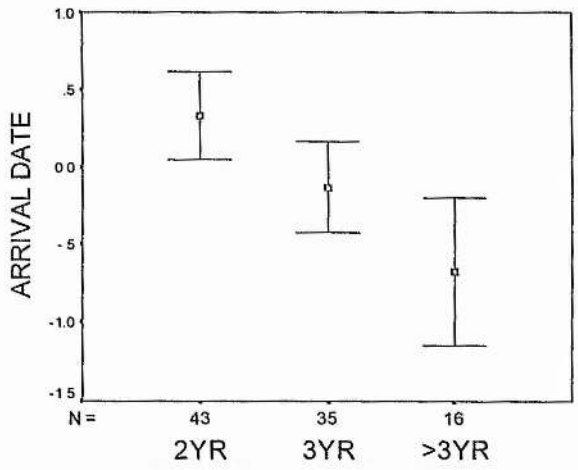


Figure 3.2. Relationship between arrival date and age. Old birds arrive earlier than young birds, the largest difference being between 2nd years and birds older than 3 years

I performed a stepwise regression analysis in order to test if arrival date could be predicted from any of the five song principal components. This was done for all age classes pooled together, and separately for each age class. The rationale of this duplication of analysis is that the pooled analysis will tell us whether females can gain information about male arrival alone from the characteristics of the song. The separate analysis for each age class examines whether the covariance of song characteristics with fitness is equally relevant at different ages. This pooled analysis yielded an overall significant model ($F=4.415$, $df=5,86$, $p<0.001$;

R^2=20.4%) in which the following two variables were significant: song length (β=0.324, t=3.36, p<0.001) and song repertoire size (β=-0.202, t=-2.075, p<0.041).

The regression analyses for each of the three age classes separately can be compared between them to test whether the relationship between arrival date and song characteristics holds true for all age classes. Only for 2YR was the overall model significant (F=3.803, df=5,60, p<0.005; R^2=24.1%); neither of the other two groups reached significance: 3YR: (F=1.105, df=5,37, p<0.375; R^2=13.1%); and >3YR: (F=1.839, df=5,16, p=0.162; R^2=36.4%). However, this is probably caused by the different sample sizes available for the different groups, and an analysis of the slopes for each of the components can be more revealing of age-related patterns (Table 3.2). Two general trends can be seen. On the one hand, the effect of song length seems to hold for all age classes. However, that of repertoire size seems to be only important for 2YR, disappearing or even tending to have an opposite effect later in life. Overall, the variable that best explains arrival date is a different one for each age class (song length for 2YR, repertoire fixation for 3YR, and song versatility for >3YR).

	2YR			3YR			>3YR		
	β	t	p	β	t	p	β	t	p
Song repertoire size	-0.27	-2.35	0.022	-0.15	-0.98	0.335	0.36	1.65	0.119
Song length	0.38	3.34	0.001	0.15	0.94	0.352	0.48	1.97	0.066
Song versatility	0.06	0.48	0.631	-0.14	-0.86	0.397	0.52	2.47	0.025
Element rate	0.08	0.73	0.466	0.04	0.24	0.815	-0.09	-0.38	0.706
Repertoire fixation	-0.07	-0.58	0.562	-0.31	-1.95	0.059	0.19	0.86	0.402

Table 3.2. Slopes and significance of the five song characteristics in predicting arrival date in the three different age classes.

Longitudinal analysis

The analysis of the differences between the same birds in their 2nd and 3rd years confirmed the results of the cross-sectional analysis (see Table 3.3), in that birds in their third year had larger song repertoires than they had in their second year. The differences in song versatility were not significant, although the trend was in the same direction as for the results obtained in the cross-sectional analysis. Another significant difference which was not found in the cross-sectional analysis was that of element rate: birds increased their element rate in their third year.

	second year	third year	t	P
Song repertoire size	-0.362 (1.19)	0.305 (0.73)	-2.742	0.011
Song length	-0.181 (1.17)	0.001 (1.10)	-0.674	0.507
Song versatility	-0.108 (0.79)	0.186 (1.01)	-1.730	0.096
Element rate	-0.371 (1.08)	0.105 (0.78)	-2.340	0.028
Repertoire fixation	-0.181 (0.81)	0.274 (1.14)	-1.584	0.126

Table 3.3. Results of the longitudinal comparison between birds in their 2nd and 3rd year for the five song principal components. Data are means, with standard deviations between brackets. The tests were paired t tests (all df=24).

I tested whether there was a correlation between the increase in repertoire size between the two years and the change in the rest of the components. Only one correlation approached significance, that of element rate ($r=-0.377$, $n=25$, $p=0.063$), suggesting that increases in repertoire are related to reductions in element rate. The prediction that birds might achieve increases in repertoire size by increasing repertoire fixation (Chapter 1) was not supported ($r=0.063$, $n=25$, $p<0.50$).

Birds did not arrive earlier in their third year than their second year (paired t test: $t=1.305$, $df=20$, $p=0.20$). However, there was a tendency for changes in arrival date between years to be negatively correlated with changes in song length ($r=-0.429$, $n=21$, $p<0.052$). In other words, birds that arrived earlier in their 3rd year than in the 2nd one, tended to decrease their song length, whereas those that arrived later, tended to increase it. No other change in song characteristics was correlated with change in arrival date (all tests: $-0.04<r<0.18$, $p>0.40$).

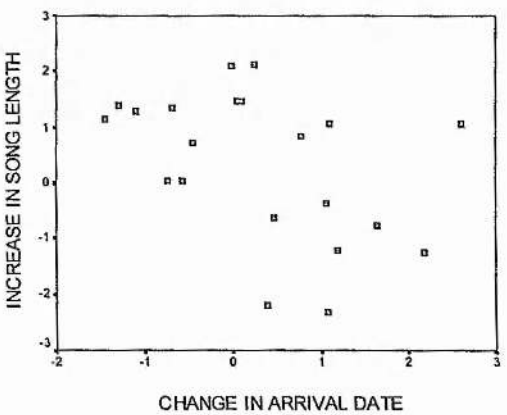


Figure 3.2. Birds that arrived earlier in the 3rd year than in the 2nd one decreased their song length, whereas those that arrived later, increased it ($r=-0.429$, $n=21$, $p=0.052$). Change in arrival is measured in the sense that positive values indicate an earlier arrival in the second year.

Discussion

The results show that repertoire size, song versatility and element rate increase with age in the willow warbler. A correlation between age and versatility had already been described for this species (Järvi 1983), although based in a very limited sample. This pattern was not a primary consequence of different mortality, although this factor was not tested specifically here (see Chapter 5). The longitudinal and the cross-sectional analyses came to similar conclusions, although the increase in element rate was not apparent in the cross-sectional analysis. This may be due to the greater sensitivity of the longitudinal approach for detecting changes.

Individual birds increased their song repertoire in their 3rd year compared with that in their 2nd. This phenomenon has been found in other passerines, like pied flycatchers *Ficedula hypoleuca* (Espmark & Lampe 1993) or canaries *Serinus canaria* (Nottebohm & Nottebohm 1978), and it could be taken to indicate that birds are capable of learning new songs late in life. However, laboratory and field experiments indicate that birds do memorise far more songs than they produce (e.g. Marler & Peters 1981; Marler & Peters 1982; Nelson 1992). Thus, it is possible that the new elements produced in the 3rd year were memorised in their first year, and had not been sung before. The increase in song versatility between the 2nd and 3rd years indicates that not only the content, but also the organisation of the song repertoire changes with age. Males that have been neighbours for several years tend to share elements to a greater extent than new neighbours or birds at greater distances, and even the order of elements is copied (author's unpublished data). This suggests that the plasticity of the song system in willow warblers might be considerable.

There was a tendency for increases in repertoire size between years to be negatively correlated with increases in song length. This result could be due to a trade-off between two main characteristics of song, namely song repertoire and performance. If a larger repertoire is more effective at territory defence and mate attraction than a small one, birds with larger repertoires might trade off this characteristic by reducing the average song length. However, this result runs counter the predictions of the monotony-threshold hypothesis, according to which birds with more monotonous repertoires would adopt less continuous singing styles in order to reduce habituation (Hartshorne 1973). In Chapter 1 I hypothesised that birds might optimise their repertoire organisation with age, and that increases in repertoire size might come about by increasing the fixation of the repertoire, i.e. by increasing the homogeneity of each song cluster (see also Podos et al. 1992). This

prediction was not upheld in the analysis, nor was there any evidence that repertoire fixation changes with age.

The relationship between arrival date, age and song characteristics is a complex one. Although the cross-sectional analysis showed that birds arrived earlier the older they were, this pattern was not significant in the longitudinal analysis. This might be due to an increased survival of early arriving males (see Chapter 5). Early arrival can be an indication of genetic quality (Møller 1994), to the extent that migratory behaviour is under genetic control (Berthold 1993) and that the very frequent fluctuations in the conditions of the wintering grounds can be an important source of additive genetic variance (Fitzpatrick 1994). However, the low repeatability of arrival date (Chapter 4) suggests that this trait is probably condition-dependent. Song characteristics were found to be a good predictor of arrival date. This could allow females to use song as a cue to quality for pairing or extra-pair partners (see Chapters 6 and 7). The relation of song characteristics to arrival date was not stronger for young birds than older ones. Rather, the characteristics that predicted arrival date were different at different ages, except song length, which showed a similar relationship with arrival date throughout the three age classes. In the case of the song repertoire, one might expect that the costs of learning a repertoire for second years should outweigh those of maintaining and updating it incurred by older birds. The results fit this pattern to some extent in the sense that in older birds, characteristics of song organisation (repertoire fixation for 3YR, versatility for >3YR) are more dependent on arrival date than repertoire size. A similar argument has been made in the case of pheasants *Phasianus colchicus* (Grahn & von Schantz 1994), where spur length is a sexually selected character. In this species, spur length was less informative in terms of phenotypic quality in young birds than in older ones. This might be so because young males present delayed development of this trait, and thus incur in lower costs than older ones. Although the results presented here may to some extent be dependent on the different sample sizes available for the three age-classes, they still suggest that selective pressures acting in song repertoires may be different at different ages.

Chapter 4

REPEATABILITY AND HERITABILITY OF SONG CHARACTERISTICS

Abstract

Between-year repeatabilities for several song characteristics and arrival date were measured in a population of willow warblers. The results show that repertoire size has a significant repeatability (0.497), and that the song characteristic with the highest repeatability was song versatility (0.770). An analysis of the evolvability of these traits suggests that, given a minimum heritability, repertoire size would have higher residual and additive genetic variance than song versatility. Father-son regressions for all song characteristics were not significant due to the small sample size and the low heritability that is expected from the repeatability results.

Introduction

Models of sexual selection propose that exaggerated ornaments in males are maintained either as a result of a genetic covariance with the female preference (Fisherian model), or alternatively because, by imposing a handicap on bearers, they become an honest signal of quality (Andersson 1994). One of the main assumptions of all models is that secondary sexual traits are heritable in males. If this is not the case, there can be no indirect selection on female mating preferences.

The evolvability of a trait, rather than its heritability, has been recently advocated as the appropriate measure of the amount of variability present in a trait (Houle 1992; Pomiankowski & Møller 1995). Evolvability is the coefficient of additive genetic variance (Houle 1992), and becomes specially useful when predictions about the response of a trait to selection, or about the forces that maintain its genetic variability, are to be made. This is because heritability is the ratio of additive genetic variance to phenotypic variance ($h^2=V_A/V_P$), and phenotypic variance in turn is the sum of residual and additive genetic

variance ($V_P=V_A+V_R$). Thus, a low heritability can be the result of either a low additive genetic variance or high residual variance (Houle 1992). A recent review shows that sexual traits do indeed have high levels of additive genetic variance, regardless of the effects of sexual selection (Pomiankowski & Møller 1995)

Birdsong is a secondary sexual trait as, in the great majority of species, at least in temperate regions, only males sing. Many species present extreme elaboration of this trait, with large and complex repertoires. Song repertoires have been argued to be related to sexual selection in many species, both in the modes of male-male competition and female choice (Searcy & Andersson 1986). Although female preference for song repertoires is commonly found in laboratory experiments, evidence for link in the field is very limited, and most studies find repertoire size to be a poor predictor of mating success (Searcy 1992). It is not clear which model of sexual selection explains best the evolution of song repertoires. A Fisherian hypothesis has been advocated by Catchpole (1980) in the case of *Acrocephalus* warblers. Also, in a broader comparative study, the strength of female preference was found to be correlated with the average repertoire size of the species (Searcy 1992). This is one of the predictions of the Fisherian coevolutionary hypothesis, but it could also be explained from a good-genes standpoint, if there is between-species variation in the variance of viability. Evidence for good-genes effects is more scant, but a correlation between viability and repertoire size has been found in great reed warblers *Acrocephalus arundinaceus* (Hasselquist et al. 1996).

These conflicting patterns could be clarified if we knew the indirect benefits that females would accrue by mating with males with large song repertoires. However, to the best of my knowledge, no estimate of the heritability of song repertoire size in any bird species has ever been published. A genetic basis for this trait is suggested by the existence of geographical variation in the repertoire size of the marsh wren *Cistophorus palustris* (Kroodsma & Canady 1985). Differences in repertoire size between the different strains of domesticated canaries *Serinus canaria* (Mundinger 1995) suggest a possibly substantial heritability of the trait, although no controlled selection experiment has measured the effect. However, it is not known whether there is heritability for repertoire size in any natural bird population. Although knowing the heritability of this ornament would not solve the controversy between the different models, it would certainly help to guide future research in the field.

Heritability estimates for song repertories are difficult to obtain. Experiments in captivity are problematic because the song tutoring regime does not usually lead to the full variance

found in nature (e.g. Kroodsma & Canady 1985). Field studies on the other hand require very big sample sizes, given the low survival rate of offspring, and results can be confounded if there are high levels of extra-pair paternity. The only aspect of song in birds for which heritability is known is song rate in the zebra finch *Taeniopygia guttata* (Houtman 1992). However, although song rate is an important trait in mate choice in many species, including the willow warbler (Radesäter et al. 1987), it is effectively a performance trait, rather than an ornament.

Repeatability (or intraclass correlation coefficient; Sokal & Rohlf 1995) is a measure of the proportion of phenotypic variance of a trait which is due to consistent differences between individuals A high repeatability value indicates that the character being measured has a high stereotypy in each individual, and that differences between individuals are high enough for selection to act on that character (Boake 1989). An important property of repeatability is that it sets an upper limit to heritability (Falconer & Mackay 1996). However, a measure of repeatability is highly dependent on the time scale over which it is taken. This time scale should be relevant to the organism. In the case of animals which can breed for several years, between-year rather than same-year repeatabilities seem the most sensible approach.

I have estimated the repeatability of several song characteristics in the song of the willow warbler. These figures give an indication of the maximum heritability that can be expected. I have also performed parent-offspring regressions in order to estimate the heritability of these different song characteristics. This is not an ideal method, and the small sample size and the existence of extra-pair paternity reduce the reliability of the results. I am aware of the limitations of this method, but my approach here is one of exploration. Arnold (1994) calls this method the "genetic without tears" approach, and finds it justifiable if "the behavior in question has never been scrutinized from the heritability point of view" (p. 33). This is the case for song repertoire size in birds.

Methods

Adult birds were caught in mist nets and marked with unique colour ring combinations so they could be identified in the field. Around 50% of adult males returned to breed a second year, and I attempted to record the song of as many of these males as possible in several years. Nests were searched for from the moment of pair formation, as females start building shortly after they pair. Most nests were found during the nestling stage, by following both parents while they were feeding the young. Nestlings were ringed with the

same combination of colour rings in each clutch. Furthermore, all recruits were eventually caught after territory settlement, so the precise identity of each bird was always obtained.

Song characteristics were measured for each male from recordings of 15 consecutive songs, recorded at least 5 days after arrival and before pairing. Five independent principal components were extracted from a larger set of variables measured in the song recording (Chapter 1). These principal components were 1) song repertoire size; 2) song length; 3) song versatility; 4) element rate; and 5) repertoire fixation. Arrival date is defined as the date a male was seen defending a territory for the first time. As the variance in arrival date changed in the different years, data for this variable are standardised per year, in order to be able to pool data together.

Repeatabilities were calculated taking into account the warnings of Lessells & Boag (1987). I divided the data set in two groups: a first one of 25 individuals recorded in their second and third years; and a second set of males recorded in several years after their third year (6 individuals recorded in 2 years and 4 individuals recorded in 3 years). This is because song characteristics change between birds in their second and third years (Chapter 3), and it can assumed that the trait is not fully developed before that age. Repeatabilities were calculated from the average mean squares values between and within individuals. As an indication of the significance of the repeatability analysis, I used the probability values associated with the F value of the ANOVA with which the mean squares were calculated (Butlin & Hewitt 1986).

As all principal components scores have a variance of 1 and a mean of 0 by definition, their potential coefficients of additive genetic variance could not be calculated (Houle 1992). However, for some principal components, the major contribution to variability comes from a single variable, and this can be used instead for comparative purposes. This was the case of song repertoire (correlation with element repertoire =0.885), and song versatility (correlation with average song versatility =0.867). The coefficient of additive genetic variance was calculated following Houle (1992): $CV_A=100*(\sqrt{V_A}/\text{mean of the trait})$, where V_A is the additive variance of the trait.

The heritability analysis was done by performing father-son regressions of the song characteristics and arrival date. None of the regressions was significant, and thus the heritability coefficients ($h^2=2*\beta$)(Falconer & Mackay 1996) were not calculated. This analysis is very limited for several reasons. On the one hand we have the sheer limitations of the sample size: very few birds ringed as nestlings survive and return to settle on a

territory. Out of these, for fewer still (n=16) are there good recordings for both father and offspring. The second problem is that there is a substantial level of extra-pair paternity in the population (18% of nestlings), which makes assumptions about putative paternity unreliable. For a limited number of individuals paternity data are known with certainty (Chapter 7). This allowed the rejection of 3 individuals from the regression analysis, as the putative social father was excluded. Unfortunately, I did not have recordings of the most likely father of these individuals. In the other 4 individuals for which paternity data were available, the social father was not excluded from being the real father (exclusion probability = 97%). For the rest of birds (n=9), no paternity data could be obtained, and they were not excluded from the regression analysis.

Results

The results of the repeatability analyses (Table 4.1) show that the repeatabilities for all song characteristics increased with age. In other words, birds have more stable and repeatable songs after their third year. The highest repeatability was that of song versatility. Repertoire size had lower repeatability than song length or versatility, but it was also significant. Arrival date had a low repeatability.

	REPEATABILITIES	
	2YR - 3YR	**3YR - >3YR**
Song repertoire size	0.143	0.493*
Song length	0.333	0.527*
Song versatility	0.410	0.770**
Element rate	0.048	0.295
Repertoire fixation	-0.530	0.386*
Arrival date	0.237	0.163

Table 4.1. Repeatabilities for the five principal components of song characteristics and arrival date. Significant values are marked with asterisks and refer to the F value of the ANOVA. Degrees of freedom were 8,16 for the 3YR->3YR analysis and 23,25 for the 2YR-3YR.

The results confirm the conclusions of the analysis of changes in song with age, in that songs seem not to change much after the third year (Chapter 3). The repeatability analysis of the birds older than 3YR seems to be the most meaningful one, as birds have fully

developed songs. I also calculated the repeatabilities of the direct variables which contribute to most strongly to the variation in two principal components: repertoire size (direct variable: element repertoire) and song versatility (direct variable: average versatility). The repeatabilities of these variables are very similar to that of the principal component to which they contribute (element repertoire: 0.504; average versatility: 0.754). Assuming that the heritability has to be lower than these figures, I have calculated the components of variance for a range of potential heritabilities (Table 4.2). An interesting pattern that emerges is that, for the same heritability, additive genetic variance for element repertoire is higher than for song versatility. This is due to the higher phenotypic variance of this trait (see also Chapter 1).

	element repertoire			**average song versatility**			
potential heritability	**0.1**	**0.3**	**0.4**	**0.1**	**0.3**	**0.4**	**0.7**
CV_A	9.8	15.8	18.2	3.8	6.6	7.7	10.1
CV_P	28.8	28.8	28.8	12.1	12.1	12.1	12.1
CV_R	27.4	24.1	22.3	11.5	10.1	9.4	6.6

Table 4.3. Coefficients of additive, phenotypic and residual variance for a range of potential values of heritability for two traits: element repertoire and average versatility.

Father-son regressions were non-significant for all characteristics (Table 4.3), so the null hypothesis that $\beta=0$ cannot be rejected. Plots of the father-son data are presented for song repertoire (Fig 4.1a) and song versatility (Fig 4.1b).

	β	t	p
Song repertoire size	0.110	-0.200	0.845
Song length	-0.120	-0.585	0.570
Song versatility	0.403	-0.872	0.400
Element rate	0.270	-0.700	0.499
Repertoire fixation	-0.589	-1.306	0.218
Arrival date	-0.221	-1.367	0.197

Table 4.3. Results of the father-son regression for the five principal components of song characteristics and arrival date. Sample size was 12 for the song characteristics and 13 for arrival date.

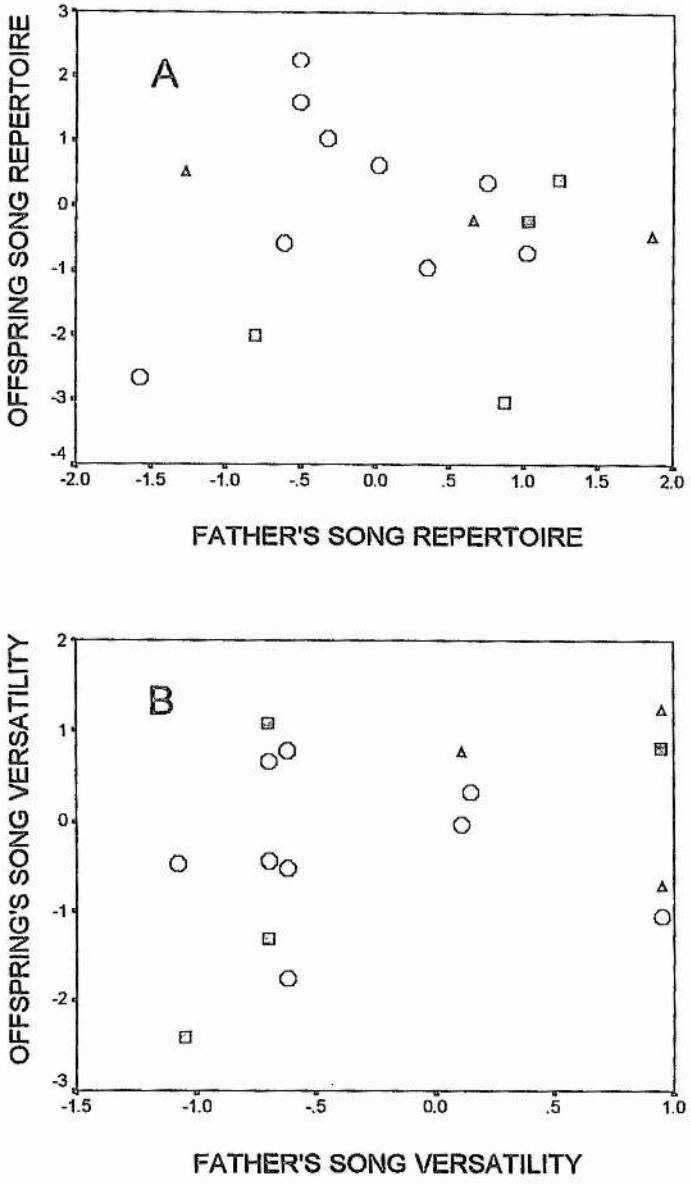


Figure 4.1. Plots of offspring's song repertoire values (A) and song versatility (B) against that of the father. Paternity is known for the cases represented by black squares. Red triangles are extra-pair young and blue circles birds whose paternity was only putative.

Discussion

The repeatability values found for repertoire size suggest that its heritability must be lower than 0.49 in this population of willow warblers. The relationship between repeatability and heritability depends on the environmental component of genetic variance. If the latter is not large, repeatability and heritability will be similar (Boake 1989). It is not at all surprising that given this low level of repeatability, the sample size used in the father-son regression was not sufficient to yield a significant estimate. Of all song characteristics, song versatility

was the one with the highest repeatability. These repeatability values were obtained for birds measured in several years. A comparison with the repeatabilities found for the same characteristics in a within-year analysis (Chapter 1) may help us to ascertain the nature of the sources of variation. The first main difference is that most characteristics show the highest repeatabilities in the within-year analysis, which suggests that the sample size of 15 songs is a good representative of the bird's repertoire. Repeatabilities are much lower for the analysis between the second and third year, and this confirm the results of the analysis of the change in song with age, in that songs are not fully developed in the second year. Later in life, the repeatabilities become very high, which indicates that the trait is developed and there is less scope for change.

This indicates that, if repertoire size is correlated with male quality, this quality cannot be assessed in second-year birds, due to the large component of environmental variance which influence the development of the trait. The consequences of this is that the high between-year variation would not to allow females to obtain genetic benefits by choosing males on this variable, unless age can be assessed by other means (see also Alatalo & Lundberg 1986). However, if repertoire size is a condition-dependent characteristic, and condition is related to direct benefits (e.g. nest attendance, fertility...), choices based on this characteristic would be directly selected.

The song characteristic with the largest repeatability was song versatility, and this was so for the two analysis at different ages. These results indicate that individual differences in song versatility are the largest and more consistent within and between years. In terms of female choice or male-male assessment, this could indicate that, as song versatility is less sensitive to between-year variation than repertoire size, it is more likely to be a genetic quality indicator than repertoire size.

If we consider the components of variance of these two traits (using their respective correlates and assuming a minimum heritability), we can see that, although repertoire size comes out with a moderately high value of additive genetic variance (9.8% if h^2=0.1), the residual variance is still very high (27% if h^2=0.1). In the case of song versatility, however, a similar heritability would yield much smaller values of genetic and residual variance. These results do not invalidate the previous conclusions based on the repeatability estimates.
In a recent review, values of additive genetic variance for sexual characters were found to be higher than those found for non-sexual characters (Pomiankowski & Møller 1995). The authors argued that these results could be explained if directional selection also selects for modifiers that increase phenotypic variation per se without changing the phenotypic mean.

This suggestion is an exciting possibility but it has yet to be demonstrated empirically. Furthermore, there is the possibility that Pomiankowski & Møller's arguments are affected by a bias in the publication of heritabilities of sexual characters (Alatalo et al. 1997). This is because, as there is a theoretical paradigm that predicts these results, studies with small sample sizes are more likely to be published if they report high heritabilities. Notwithstanding this possibility, some implications for the traits studied here can be derived from Pomiankowski & Moller's (1995) arguments. If the predictions based on modifier theories are right, the high values for residual and additive variance in song repertoire would suggest that the trait is subject to directional selection. On the other hand, the low variances of song versatility point to a possible effect of stabilising selection.

Arrival date has been shown to be an important component of fitness in migrant passerines (Verhulst et al. 1995; Svensson 1997). The low repeatability of this trait fits with the prediction that the heritability of traits correlated with fitness should be close to zero (Fisher 1930; Gustafsson 1986).

Chapter 5

DO MALE SONG CHARACTERISTICS PREDICT SURVIVAL AND VIABILITY?

Abstract

The main distinction between Fisherian and good-genes models of sexual selection is that according to the latter, there should be a positive correlation between the male trait and the survival of the bearer and of its offspring. I found that male repertoire size correlated with number of fledglings in a population of willow warblers. Although the number of recruits was not correlated with the father's repertoire size, males with at least one recruit had larger repertoires than those without any recruit. Survival of males in their second year was negatively correlated with song length. However, among birds older than second years, element rate and repertoire size correlated positively with survival. The results are consistent with a good genes hypothesis, but other possibilities like direct benefits from males with large repertoires or differential maternal effects cannot be rejected.

Introduction

Good-genes models of sexual selection predict that viability should be positively related to the expression of sexual ornaments (Zahavi 1975; Kirkpatrick & Ryan 1991). This prediction is the fundamental distinction from Fisherian models. According to this hypothesis, only males of high genetic quality are able to express traits with a high degree of ornamentation, and females that mate which such males will produce offspring of higher viability. Genetic models seem to be able to accommodate this possibility (e.g. Grafen 1990), and the required heritability of characters associated with fitness also seems to be there (Pomiankowski & Møller 1995), although the mechanisms by which additive variance is maintained are as yet unclear. Several empirical studies have found evidence of a correlation between the size of a sexual ornament and survival for adult males (Alatalo et al. 1991; Møller 1991), and also between the trait of the father and the viability of offspring (Norris 1993; Møller 1994; Petrie 1994; Hasselquist et al. 1996; Sheldon et al. 1997).

One of the most conspicuous sexual characters in birds is that of song. Many species use complex repertoires in the context of mate attraction, and repertoires are more effective than single songs at eliciting sexual behaviour from females (Searcy 1992). Repertoire size has been found to increase with age in several species (e.g. Chapter 3). To the extent that longevity is heritable, this would represent an honest signal of genetic quality. A positive correlation between repertoire size, survival and lifetime reproductive success has been found in the great tit *Parus major* (McGregor et al. 1981). However, survival was negatively correlated with repertoire size and song versatility in a Swedish population of pied flycatchers *Ficedula hypoleuca* (Lampe & Espmark 1994). Higher viability of the offspring of males with large repertoires has been found for great reed warblers *Acrocephalus arundinaceus* (Hasselquist et al. 1996). There is evidence of neural costs associated with large song repertoires (Canady et al. 1984; Devoogd et al. 1993), which suggests that the expression of this trait is condition-dependent, and thus not amenable to cheating. This implies that variation in repertoire size signals some aspect of male quality. The present chapter examines how several characteristics of the song of the willow warbler predict male survival, number of fledglings and offspring viability.

Methods

General field methods have been described in previous chapters. Given the very high philopatry of this species, any adult bird which did not hold a territory in the area the year after it was last recorded was taken to be dead. A few individuals (4 in the 4 years) were subsequently caught in nets during routine ringing and never settled in a territory. Although a few more might have been missed, this in unlikely to affect the results.

Survival is a binary variable and was thus analysed by logistic regression. If a bird had been recorded in several years, only one data point was chosen at random. To visualise the selective regime of the different variables, I employed the cubic spline method developed by Schluter (1988), using a program written by him (GLMS; available at: http://www.zoology.ubc.ca/~schluter/splines.html). This is a nonparametric regression method designed to plot fitness functions, that is, the relationship between a fitness measure and a quantitative trait.

Results

In a stepwise regression analysis between song characteristics and number of fledglings, repertoire size was the only song characteristic that was not excluded from the model, and

it had a significant positive relationship with number of fledglings (overall regression: $F=10.278$, $df=1,23$, $p<0.01$; $R^2=0.32$; repertoire size: $\beta=0.643$, $t=3.206$, $p=0.004$; Fig 5.1). Arrival date is also a predictor of number of fledglings ($F=8.235$, $df=1,20$, $p<0.01$; $R^2=0.29$; arrival date: $\beta=-0.530$, $t=-2.870$, $p=0.009$). Given that repertoire size and arrival date are correlated, the relation between repertoire size and number of fledglings could be due to the common correlation with arrival date. I performed a Kendall's partial correlation between repertoire size and number of fledglings, controlling for the effect of arrival. The partial correlation shows a genuine effect of repertoire size in number of fledglings ($\tau=0.472$, $n=23$, $p=0.0022$). Number of fledglings is probably highly correlated with clutch size, given that no nest mortality was recorded in this population in the course of the study.

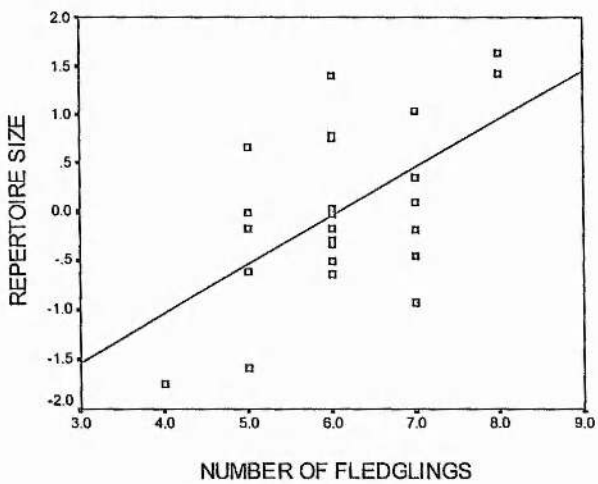


Figure 5.1. Males with a larger repertoire size fledge more young. None of the other song characteristics predicted number of fledglings.

Number of recruits could not be predicted by any song characteristic (multiple regression: $F=0.442$, $df=5,18$, $p=0.813$; all variables $p>0.20$). As most of the nests contributed either one or no recruits, I performed a stepwise logistic regression in which the dependent variable was either none or at least one recruit. The final model is significant overall ($\chi^2=5.997$, $df=1$, $p=0.0143$), and included a constant ($B=1.73$, $df=1$, $p=0.002$) and song repertoire, whose effect was close to significance ($\beta=2.30$, $df=1$, $p=0.082$). This indicates that the survival of at least one recruit can be significantly predicted from the song characteristics, and that song repertoire is the best predictor for it (Fig. 5.2).

The pattern of survival was different for second-years and older birds. In a stepwise logistic regression, song length was the variable which best explained survival for second years (overall model: $\chi^2=5.025$, $df=1$, $p=0.025$; song length: $\beta=-0.549$, $df=1$, $p=0.038$). In other words, second-year birds with long songs were selected against. In older birds, however,

two different variables were included in the final model: repertoire size and element rate (overall model: χ^2=7.303, df=2, p=0.026; repertoire size: β=0.694, df=1, p=0.087; element rate: β=0.754, df=1, p=0.031): older birds with high element rates and large repertoires were more likely to survive. A further test of the result that birds with short songs are selected in the first age cohort can be made by looking at the frequency distribution of the variable in the two age cohorts, i.e. before and after selection (Fig 5.3). Note that the two means were not statistically different (Chapter 3).

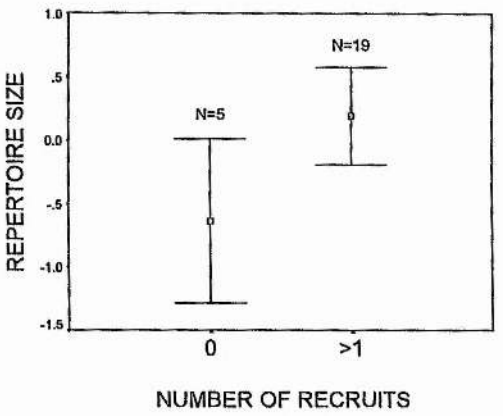


Figure 5.2. Males which produced no recruits had marginally smaller repertoires than those which had at least one recruit in the population.

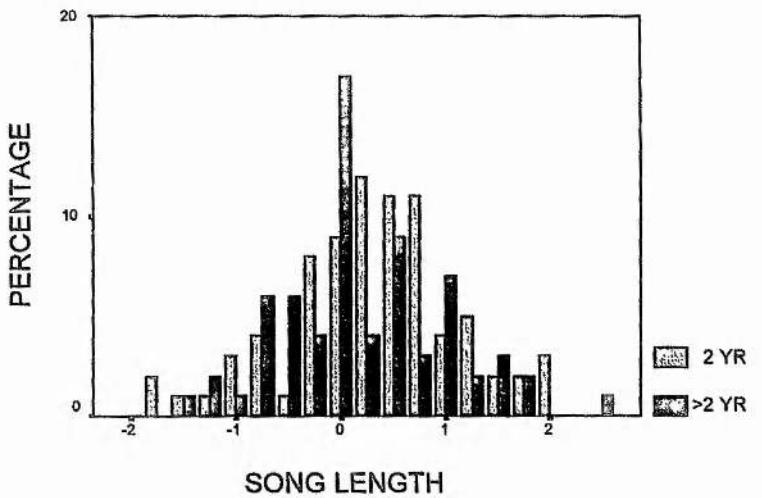
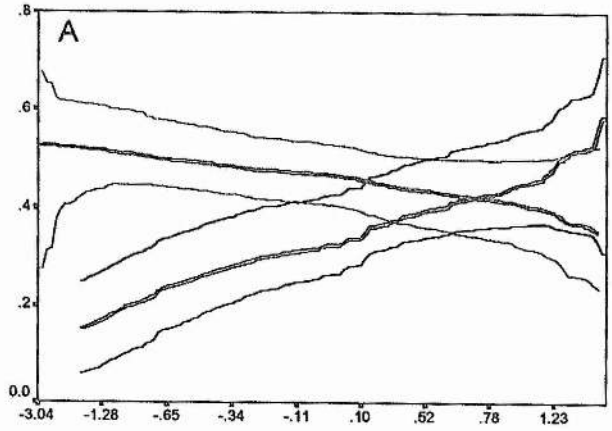
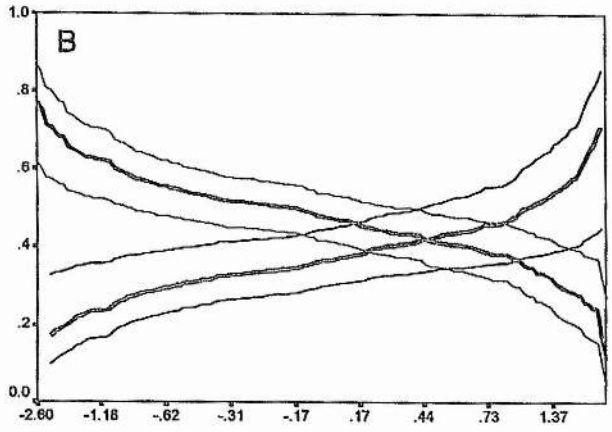


Figure 5.3. Frequency distribution of song length for second years (blue) and older birds (red). The two distributions are statistically different (G=31.2, df=6, p<0.001).

The three variables correlated with survival were not the same in the two age classes. Cubic splines showing the effect of the three variables considered above are shown in Fig. 5.4a-c. Interestingly, the effect of the variables tend to be inverted: second years with high values in the three variables tend to be selected against; whereas the effect is positive for older birds.





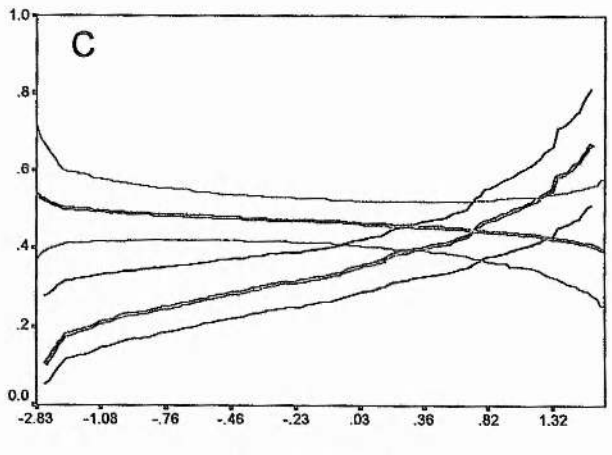


Figure 5.4. Fitness functions relating probability of survival (y axis) to three different song characteristics, as a function of age (second years: red; older birds: black): A) element repertoire; B) song length; and C) element rate. The thick lines are the fitted cubic splines and the thin lines at each side are ±1SE (obtained through 1000 bootstraps).

Discussion

The results show that there is a correlation between a male's song repertoire size and the number of fledglings from his nest, as well as with the probability he contributes at least one recruit in the population. These results are consistent with a good-genes hypothesis, which would predict that if repertoires are honest indicators of male genetic quality, females paired to these males would obtain genetic benefits (Kirkpatrick & Ryan 1991). However, explanations based on direct benefits or on differential female investment are also consistent with the results. Male willow warblers normally feed and care for the young throughout the nestling period and after fledging, and it is possible that the increased viability is due to an effect of better territories or paternal care. Another possibility is that females that pair with males with large repertoires are of higher phenotypic quality, which has been shown for barn swallows *Hirundo rustica* (Møller 1991). This is the most likely explanation for the correlation between number of fledglings and repertoire size, specially as number of fledglings is likely to be a close correlate of clutch size. An alternative explanation is that females paired to males with large repertoires increase the clutch size and/or the investment in the clutch. This second hypothesis would correspond to a good-genes hypothesis, but unfortunately the correlative nature of this study cannot tease apart these two possibilities.

Maternal or direct benefits effects tend to be treated in studies of sexual selection as rather fastidious confounding variables, which are normally cast aside with vague arguments (e.g. Møller 1994). However, that females invest more in the offspring of attractive males has already been shown in several species (Burley 1988; De Lope & Møller 1993), and this could be predicted from arbitrary trait models (Kirkpatrick & Ryan 1991). Experiments are necessary to tease apart the effects of paternally inherited genes and maternal effects (i.e. any environmental source of variance provided by relatives: Cheverud & Moore 1994). A good example of this experimental approach was that of Petrie (1994) with peacocks *Pavo cristatus*, where females were experimentally allocated to males of different tail ornamentation and the offspring fostered by domestic hens. The results showed increased growth and survival of chicks fathered by long-tailed males. A similar situation was achieved in the wild by Sheldon et al. (1997), by looking at the growth of extra-pair paternity chicks compared to their half-siblings in the collared flycatcher *Ficedula albicollis*. Here again the body condition of chicks was related to the size of the father's ornament. Both these studies, however, assume that females are not investing differently in the eggs fathered by the different males. That they may do so is suggested by the discovery that

avian eggs contain testosterone (Schwabl 1993). The amount of testosterone (T) varies from species to species, and between the eggs of the same clutch. Yolk T correlates with the values of circulating T in the mother at the time of yolk production (Schwabl 1996) and influences the development of the chick, via a higher begging rate and aggressiveness (Schwabl 1996). Chicks that hatch from eggs that have higher amounts of T, or that have had their yolks experimentally injected with T, also obtain higher ranks in the social hierarchy after fledging (Schwabl 1993). Some of the results of the studies that claim good gene effects in the offspring, like those of Sheldon et al. (1997) or Petrie (1994), could be explained in terms of differential allocation of T by females. There is no need to postulate an adaptive explanation, although we know that adaptive allocations of T have been shown in some species (Schwabl et al. 1997). The mechanism I envisage only requires that females produce a higher level of T when they mate with an attractive male. As the egg that is going to be fertilised by this sperm is still being formed (Colegrave et al. 1995), and yolk being incorporated into the follicle, it follows that the increase in circulating T will be transferred to the yolk. Careful experiments should be carried out to examine this possibility.

The most interesting result concerning the survival of adults was the interaction of age with several song characteristics. Song length was the best predictor of survival in second-years: birds with long songs were selected against. In older birds, however, two different song components, element rate and repertoire size, were positively correlated with survival. This pattern can be examined by looking at the cubic splines of these three components for the two ages (Fig 5.4). Although song length is not a significant predictor of survival in old birds, or repertoire size in second-years, the slopes suggest that the effect of these two variables in survival is opposite for the two age cohorts. What is this pattern due to?

It has been argued that, when the development of the trait is different between young and older birds, only old males, which have developed the trait to its full extent, will be paying the handicap price of it. Veiga (1993) manipulated the size of the badges of house sparrows *Passer domesticus* and found that males with enlarged badges acquired more nest sites. The survival consequences of the experimental manipulation showed an interaction with age (Veiga 1995): enlarged badges decreased the chances of surviving in young males, but there was a trend for the opposite to be the true in older males. Veiga argued that increased breeding success by young males was checked by the higher costs of breeding that these males bore. Another similar case is that of the spur length of male pheasants *Phasianus colchicus,* which correlates with age (von Schantz et al. 1989). This trait is a good predictor of survival in old individuals, but not in young ones (Grahn 1993). A correlation between spur length and a measure of fitness (fluctuation asymmetry) was only significant in old males,

suggesting that only in this age class did spur length indicate male quality. Mate choice was consistent with this pattern: females chose long-spurred males only in the case of old males, and the mating success among young males was not related to this trait (Grahn & von Schantz 1994). The explanation offered for this set of results was based on the age-dependent development of the spur length: Grahn & von Schantz argued that young males are paying a lower cost in growing these ornaments. As a result of this delayed development, the handicap associated with it may not be apparent before the spurs reach their full length. However, this argument could only hold true if the variance and the range of the trait were different between the different age classes. Still, it is difficult to see how a negative correlation between the trait and survival could follow from this argument.

In the case of the song characteristics studied here, element rate and repertoire size do increase with age, but song length does not (Chapter 3). The variance of these characteristics between age classes are not significantly different (Levene's tests for equality of variances, all $F<2.1$, $n=108$, $p>0.25$) and the ranges overlap to a great extent. Why then are second-years not subject to the same handicapping effects observable in older birds? A failure to show a relationship is not proof of its non-existence, so it could be argued that the different selection regimes between age classes that can be inferred from Fig 5.4 are an artefact. Fair enough, but what if they are not an artefact?

The fact that song length is selected against in second years can be explained if low quality males are trading off survival for current reproduction (Williams 1966; Orell et al. 1994). The development of long songs may result in an higher reproductive success. Although females do not seem to choose males on the basis of this variable (Chapter 6), males with longer songs were less likely to be cuckolded (Chapter 7). Males with long songs might also give more paternal care that would explain the increased mortality. If the reproductive success of low quality males increases in their first breeding year as a result of singing long songs, this could constitute an adaptive strategy, regardless of the reduction in the probability of survival. Male quality need not be genetic, and the strategy could work as well if survival is condition-dependent. The positive correlation between survival and element rate and repertoire in older males would reflect the expected correlation between trait expression and male quality after the first episode of selection has reduced the variance in genetic quality. These traits do not change after birds are older than three-years (Chapter 3), and thus there would be no room for developmental trade-offs.

To conclude, males with large repertoires produced more young and were more likely to have recruits in the population. This is consistent with a good-genes hypothesis, but other

hypotheses can account for the same result. Repertoire size and element rate predicted survival in birds older than second-years. The complex relation between age, survival and song characteristics could be explained by an inter-generational trade-off, if low quality males invest disproportionately in the development of attractive songs.

Chapter 6

FEMALE CHOICE AND MALE SONG CHARACTERISTICS

Abstract

Male willow warblers have song repertoires which vary in complexity along several dimensions. Song repertoires may be used by females when choosing a mate. This was tested by studying the pairing patterns of a population in relation to five song characteristics in four different years. Pairing date was negatively correlated with song repertoire size in one year, and with song versatility in another year. The variable that best explains how soon a male gets paired is male arrival date, and this is consistently significant in all years. Among the song characteristics, the best predictors of pairing date in any given year were also the most variable ones. Contrary to expectations, the predictive value of the song characteristics and arrival date for pairing date in a given year is negatively correlated with the amount of female competition in that year.

Introduction

In many songbirds, males possess more than one copy of their species-specific song. These sets of songs sung by a single male are called song repertoires. The evolutionary significance of repertoires is usually explained in terms of sexual selection, both within the context of male-male competition and that of female choice (Searcy & Andersson 1986). Female preference for large song repertoire sizes has been studied through laboratory experiments and field observations. Evidence from laboratory data is highly consistent and shows that when females are exposed to large repertoires they display higher levels of sexual arousal than when they hear small repertoires (e.g. Searcy & Marler 1981; Catchpole et al. 1986; Vallet et al. 1998). Field data however are not as straightforward. Strong directional preferences have been found in *Acrocephalus* warblers (Catchpole 1980, 1986; Buchanan & Catchpole 1997), and in the European starling *Sturnus vulgaris* (Mountjoy & Lemon 1996), but in other species where this issue has been studied, the effects of song

repertoire on female choice are either non-existent or disappear when territory quality is controlled for (Howard 1974; Krebs et al. 1978; Yasukawa et al. 1980; McGregor et al. 1981; Searcy 1984).

Searcy (1992) has suggested that the discrepancy between laboratory and field data might be due to the fact that female preference for large song repertoires is exerted more strongly in extra-pair matings than in social mate choice. Another possible explanation for this discrepancy of results was proposed by Podos et al. (1992). These authors looked at the variation between the different songtypes present in the repertoires of song sparrows, *Melospiza melodia*, and found that birds with larger repertoires had songtypes more similar between them. This implies the existence of a constraint in the design of a song repertoire. They went on to suggest that the number of song types per se might not be a proper measure of repertoire complexity, and that the variation between songtypes could be the important variable that was being selected by females (Nowicki & Podos 1993).

In the willow warbler, although it has been shown that males with high song rates pair earlier (Radesäter et al. 1987; Arvidsson & Neergard 1991), it is not known whether the strong differences between males in song complexity are also used in female choice. An early study reported a positive correlation of song repertoire size and song versatility with pairing date (Järvi 1983). However his study involved a very small sample size, used one-tailed statistical tests, and did not take into account the confounding effect of territory quality.

The organisation of the song repertoire in this species suggests that the approach of Podos et al. (1992) for analysing the song of the song sparrow might also be useful. However, the results of the song analysis (Chapter 1) suggest that the constraint in repertoire composition found in the song sparrow might not exist in this species. Instead, my approach will take into account how independent components of song characteristics predict female choice.

Good-genes theories for the evolution of male ornaments suggest that these have evolved because they are costly to produce and thus reflect a male's quality (Kirkpatrick & Ryan 1991). Some evidence in this direction has been found in the great reed warbler *Acrocephalus arundinaceus*, where repertoire size correlates with offspring viability (Hasselquist et al. 1996). In the case of the willow warbler, there is a complex pattern between survival and age, and male survival correlates negatively with song length in the case of second-years, whereas in males older than second-years, survival is positively correlated with element rate and song repertoire (Chapter 5). It has been shown before in other species that female

choice can adjust to changes on the fitness information of a signal at different ages (Grahn & von Schantz 1994). The effect of male age in female choice will be explored in this context, to look at this possible adaptive pattern in female choice. It is expected that, if the main benefits of choice in females are genetic, pairing date should correlate negatively with song length in second-years, and positively with repertoire size and element rate in the case of older males.

Another interesting topic in the study of female preference is that of multiple cues. Males often exhibit several secondary sexual characteristics that could be used by females to assess the males' resources, condition or genetic quality. These multiple cues can either be redundant in terms of the information that they convey, or instead be indicators of different aspects of the male's characteristics (Møller & Pomiankowski 1993). In this study, by looking several variables in different years, it is possible to see if the multiple cues of the song of this species play a role in mate choice.

The present study investigates whether any of the different song characteristics of the willow warbler predicts speed of pairing. Data for four different years is compared in order to examine the possible variation in the importance of these song factors.

Methods

The study area was Kippo Wood, a mixed woodland of spruce, birch and willows in Fife, Scotland. A marked population of willow warblers was studied between 1994 and 1997. Sample sizes for each of the years were: 14, 19, 22 and 21. In the pooled analysis individual birds were considered only once.

The area was visited daily from mid April and the settlement of new territories and the pairing status of all birds were recorded. Arrival date is the day when a male was first seen defending a territory. Paired males can be detected because they stop singing long bouts from exposed perches and follow the female closely. Thus, pairing date is the day when a bird was first seen with a female that later copulated with him and nested in his territory. If a bird could not be found singing on a given day, but was found to be paired the day after, the first day was taken as the date of pairing. Only those males that arrived before the first female in any year were used in the analysis. Note that the actual male or female arrival dates cannot be known, and male settlement and pairing date are taken as approximations to them.

A continuous recording of 15 songs per male was used for the song analysis. Only recordings obtained more than 5 days after arrival but before pairing were used, as the song undergoes selective attrition (Nelson 1992) in the first days after arrival (unpublished). Two sets of measurements were taken. The first set is a series of direct measurements of song variables. The second set provides a detailed analysis of the organisation of the song repertoire, which was studied by cluster analysis based on the similarity matrix of songs. A principal component analysis was conducted on each of the sets of variables. These analyses yielded a total of five principal components whose nature can be clearly understood from the factor matrixes (see Chapter 1 for details). The factors are: 1) Repertoire size; 2) Song length; 3) Song versatility; 4) Element rate; and 5) Repertoire fixation.

Statistical tests were non-parametric due to the impossibility of normalising arrival and pairing dates. In the pooled analysis, data were standardised per year.

Results

Date of pairing was negatively correlated with repertoire size in 1996 (τ=-0.331, n=22, $p<0.05$), and with song versatility in 1995 (τ=-0.589, n=19, $p<0.01$). However, these song characteristics do not show any important effect or even trend in the rest of the years. None of the song components explained the pattern of pairing consistently in the four years of the study or in the pooled data set (Table 6.1). However, arrival date showed a significant effect in all years: females paired first with males that had arrived early.

	1994	1995	1996	1997	Pooled
Repertoire size	0.285	0.143	-0.331*	-0.025	0.002
Song length	-0.061	-0.182	-0.077	0.236	0.058
Song versatility	0.110	-0.589**	0.045	0.162	0.013
Element rate	0.160	0.299	-0.095	0.146	0.121
Repertoire fixation	0.086	0.026	0.041	0.075	0.058
Arrival date	0.698**	0.711***	0.381*	0.589***	0.515***
	n = 14	n = 19	N = 22	n = 21	n = 62

Table 6.1. Kendall correlation coefficients (τ) between pairing date and the 5 song characteristics, as well as arrival date, for the four years separately and for the pooled data set (standardised per year). All correlations are not significant except those marked with an asterisk ($p<0.05$) or three ($p<0.001$).

As male and territory qualities may be correlated, it is necessary to correct for the possibility that females are choosing territories rather than males. Male arrival date is not necessarily perfectly correlated with territory quality (see Discussion), but the correlation is very likely to be a positive one, so male arrival date can conservatively be used as an approximation of territory quality. When the effect of arrival date is thus taken into account, the correlation between pairing date and song repertoire size in 1996 disappears (Kendall's partial correlation coefficient τ=-0.237, n=22, p>0.05), although that of song versatility in 1995 remains significant: (Kendall's partial correlation coefficient: τ=-0.488, n=19, p<0.05).

For 14 females, pairing data for two years is available. Out of these, 5 females paired with their previous years' male, and 9 paired with a different male. For these 9 females which had a different male in each year, only one had the previous year's male available at the time of pairing. These data suggest that most females pair again with their previous males if these are available. This could be a confounding factor, as female choice could be based in previous reproductive performance rather than in song. Although I cannot control for this effect in each of the years, given the small data set in a particular year, it is possible in the pooled analysis. The results confirm the previous finding: when only new females are considered (n=40) the only significant correlation is still that of pairing date with arrival date (τ=0.490, n=40, p<0.001; Fig. 6.1).

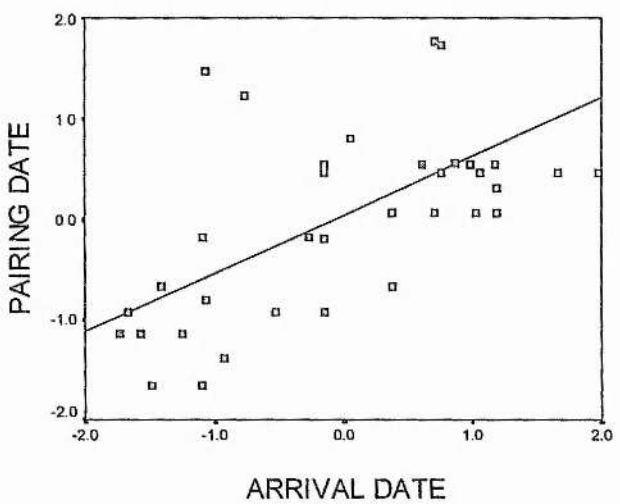


Figure 6.1. Even if only new females are considered, the better predictor of pairing date is male arrival date (τ=0.490, n=40, p<0.001). Both variables are standardised per year. Regression line for illustration only.

One of the predictions from the survival data was that, if the main benefits of choice in females are genetic, pairing date should correlate negatively with song length in second-years, and positively with repertoire size and element rate in the case of older males. In the pooled data set the correlation between pairing and song length for second-years was in the expected direction, but it was not significant (τ=0.290, n=23, p=0.056). In the case of birds older than second-years, there was no correlation between pairing and element rate and repertoire size (both τ<0.1, n=33, p>0.5). Again in both cases, arrival date was the best predictor of pairing speed (2YR: τ=0.542, n=23, p<0.001; >2YR: τ=0.498, n=33, p<0.001).

Arrival date is thus the main aspect which females select when choosing males. This could result from a concordance in habitat choice, or from a direct choice of some male characteristic which correlates with arrival date. In addition to arrival date, two song factors show in some years slight levels of female preference: song repertoire size and song versatility. It is interesting to see whether the degree of variation of these three variables in a year is correlated with their relative importance for mate choice. Previous studies have shown that when several cues are available for mate choice, females tend to use the most variable of these cues (Reid & Weatherhead 1990; Buchanan & Catchpole 1997). The three variables show significant correlations between their standard deviations and the absolute correlation coefficient which they show with pairing date in that same year, although the correlations are not significant given the small sample size (arrival date: r_s=0.400; song repertoire size: r_s=0.600; versatility: r_s=0.600; all n=4, p>0.05); (standard deviations rather than coefficients of variance are used because the values are standardised).

Competition between females has been shown to influence the stringency of mate preference. Dale et al. (1992) found that in the pied flycatcher, *Ficedula hypoleuca*, females visit fewer males before pairing when there are many females searching for mates. We can analyse how well our pairing predictors correlate with pairing speed in the different years in relation to the degree of female competition. As an index of female competition for mates, I have taken the ratio of number of females which pair in a year to the number of days in which pairing takes place. I have compared these values with the extent to which arrival date or any of the song factors (highest absolute value of the correlation coefficient between pairing date and any of the five song factors) explain the pattern of pairing. Contrary to predictions, both arrival date and song characteristics attain their highest predictive value in situations of high female competition (both correlations: r_s=0.800, n=4, p>0.05; Table 6.2).

	competition	correlation of pairing date with:	
	(females/day)	arrival date	best song predictor
1994	2.33	0.70	0.28
1995	4.75	0.71	0.59
1996	1.83	0.38	0.33
1997	1.62	0.60	0.23

Table 6.2. The amount of female competition correlates positively with the predictive value of arrival date and the best song predictor (both r_s=0.800, n=4, p>0.05).

Discussion

In two of the four years in which this study was carried out, a different song characteristic explained the choice of male to some extent. However, in the other two years, these factors did not show a similar trend. If we assume that arrival date is correlated with territory quality (see below), we should correct for the fact that females could choose territory quality rather than song characteristics. If this is done, the correlation between pairing date and song repertoire size in 1996 and song versatility in 1995 disappear. In other words, there was no overall relationship between repertoire complexity and pairing date. Other studies have found that song rate is important in mate choice in the willow warbler (Radesäter et al. 1987; Arvidsson & Neergard 1991). Unfortunately, I did not measure song rate in this study, so I am unable to assess the importance of this variable in this population. However, it is possible to conclude that song repertoire characteristics are not primarily being chosen by females in the context of social mate choice in this species.

Several studies carried out on several other species of birds show similar trends (Searcy 1992). The effect of repertoire size on pairing date in great tits *Parus major* (McGregor et al. 1981), pied flycatchers *Ficedula hypoleuca* (Alatalo et al. 1986), or song sparrows *Melospiza melodia* (Searcy 1984) is non-existent. In other species, there is some effect, but it disappears when other cues like territory quality are taken into account (e.g. Northern mockingbirds *Mimus polyglottos* (Howard 1974), red-winged blackbirds *Agelaius phoeniceus* (Yasukawa et al. 1980). The only examples where repertoire size explains pairing date to a large extent are *Acrocephalus* warblers (Catchpole 1980; Catchpole 1986; Buchanan & Catchpole 1997), and European starlings *Sturnus vulgaris* (Mountjoy & Lemon 1996).

Podos et al. (1992) suggested that the general failure of song repertoire size to explain time of pairing was due to the fact that the particular aspect of song complexity which was being preferred by females was not measured in field studies. I tested this possibility by applying the same logic to the song organisation of the willow warbler and obtaining several independent components of song characteristics (see Chapter 1). The exercise was worthwhile as, in different years, different characteristics of song contributed in different degree to pairing date. However, no overall effect could be found.

The results gave a mixed support to the predictions derived from the analysis of survival in relation to song characteristics in different age classes. Although there was a trend in the expected direction in second years (birds with long songs paired later), the predictions for older birds were nor supported. It has been shown in pheasants *Phasianus colchicus* that females only use spur length as a criterion in mate choice in the case of old males (Grahn & von Schantz 1994). This preference is adaptive, as only in that age class is spur length a male quality indicator. The results obtained here suggest that either females cannot discriminate between males of different ages, or that the benefits derived from male genetic quality are not very important.

If we put together all the studies that have been done so far, the picture that emerges shows that it is difficult to explain the maintenance of song repertoires by female choice. However, given that a very high percentage of birds show some degree of extra-pair paternity (Birkhead & Møller 1992), traditional measures of female social-male choice may not be reliable correlates of female preference. Females respond with greater intensity to large repertoires in a sexual context, and this has led Searcy (1992) to suggest that females may pay more attention to repertoire complexity in situations of extra-pair male choice. Female preference for males with large repertoires as extra-pair partners has indeed been found in the great reed warbler *Acrocephalus arundinaceus* (Hasselquist et al. 1996). Willow warblers do not stop singing after pairing (Chapter 2), and this suggests that song might be used to attract other females for EPCs. This possibility is examined in the next chapter.

The best predictor of male pairing date was date of arrival. In all four years males which arrived earlier also paired earlier (note that all males considered in the analysis were available for pairing when the first female arrived). This effect has been reported before in this species (Radesäter et al. 1987; Arvidsson & Neergard 1991), and can be a consequence of females choosing early arriving males by using some male cue which shows its arrival date, or alternatively a concordance in the choice of habitat or nest-site characteristics. The latter of these possibilities seems to be the most parsimonious one, but, is there any evidence that males choose territories on the grounds of some quality parameter?

If we look at the pattern of territory occupation between years, we can see that most birds settle in the same territory as the year before if this is available (Jakobsson 1988), and that only males which settle very late and occupy clearly suboptimal territories will change territory in the following year (author's unpublished data). That familiarity with the territory and neighbours should be important in this species is also suggested by the high degree of song sharing between birds in adjacent territories (author's unpublished data), and a between-years survival rate of about 50%, which means that the average bird returning to breed will share borders with two old neighbours. It has been shown in other species that familiarity between neighbours reduces defence costs (Eason & Hannon 1994) and can result in a higher breeding success (Beletsky & Orians 1989). There is also evidence that settling in proximity to conspecific territories can be an important cue when choosing a territory in which to settle (Stamps 1991; Muller et al. 1997).

It is not known what the characteristics of a good territory should be, so I cannot test the prediction that arrival date correlates with territory quality. Abundance of insects was not related to territory size in a Scandinavian population (Arvidsson & Klaesson 1986). A number of factors may be more important than territory size, and food need not be a limiting factor. For instance, the presence of a high song perch seems to be important for males, as these can be used to broadcast the song to a greater distance, and also might be useful for territory defence. Visibility in a territory is also an important factor that can override optimal foraging considerations (Eason 1992; Eason & Stamps 1992). Furthermore, the interests of male and female need not be the same. A territory of dense vegetation is likely to make mate guarding more difficult (Westneat & Sherman 1997), and would benefit a female seeking EPCs. A safe nesting site is obviously fundamental, but once some good nesting sites are available, greater nest site availability per se should not make a territory more valuable to a female. Nesting sites do not seem to be a limiting factor for this species in the area.

Another piece of evidence that speaks against the idea that females choose territories rather than males in the willow warbler comes from a removal experiment (Arvidsson & Neergard 1991). In this study, the first set of males to settle in an area was removed so that a second set could take the vacant territories. The comparison of the pairing dates of the second group with the settling date of the first group yielded no correlation, which points to a female choice of male characteristics rather than territory quality.

It is thus debatable that the most parsimonious explanation for the fact that early arriving males get paired earlier is that females are choosing territory quality. There are cues in the

song by which females could tell the arrival time of males. For instance, it has been shown that one of the song characteristics (not measured in this study) which females choose is song rate (Radesäter et al. 1987; Arvidsson & Neergard 1991). Song rate increases with time spent in the territory, so early arriving males will have comparatively higher song rates than late ones (Arvidsson & Neergard 1991).

However, whether females actively choose early arriving males per se or base their choice in the same habitat characteristics as males do, should not make any difference in selective terms. At the end of the day, early arriving males are being selected. This phenomenon has been termed indirect selection (Wiley & Poston 1996). Early breeding is a major component of fitness in migratory passerines (e.g. Verhulst et al. 1995; Svensson 1997), and genetic models explain how breeding date may have considerable additive genetic variance (Price et al. 1988). Thus, arrival date may be a character sexually selected by female choice in this species. Arrival date is linked to female choice in several species, although it is not normally considered as a selected character, and rather treated as a confounding variable with territory quality (e.g. Alatalo et al. 1984; Alatalo et al. 1986; Radesäter et al. 1987; Arvidsson & Neergard 1991; Bensch & Hasselquist 1992; Aebischer et al. 1996). Female choice of early-arriving males independently of other preferences has been found in the barn swallow by Møller (1994), who suggested that it might result in the acquisition of good genes for migratory behaviour. Fitzpatrick (1994) proposed that in migratory birds one of the most important sources of additive genetic variance could be the quality of their wintering grounds. Her argument is based in the fact that the direction of migration has a very strong genetic basis, and that the quality of wintering areas shows a high rate of change due to frequent fluctuations in climatic conditions. Secondary sexual characters could thus be honest indicators of the quality of the wintering areas.

There are three hypotheses behind the existence of multiple sexual ornaments in males, (Møller & Pomiankowski 1993). According to these authors, multiple male signals could be signalling: 1) multiple messages: each signal reflects a single property of the animal which is important for female choice; 2) redundant information: every signal gives an indication of an overall aspect of the animal which is chosen by the female; or 3) unreliable signals: some signals may no longer signal a given preference or benefit due to fixation or exhaustion of additive genetic variance. Research in several species suggests that multiple cues reflect different aspects of male or territory quality (Reid & Weatherhead 1990; Dale & Slagsvold 1996; Omland 1996; Buchanan & Catchpole 1997), and that multiple signals used in female choice do not always correlate with each other. Some studies have looked at mate choice by multiple cues in several years and found that females tend to use those signals that are most

variable in a given year (Reid & Weatherhead 1990; Dale & Slagsvold 1996; Buchanan & Catchpole 1997). My data fit with these results in that the different variables studied explained better the pattern of pairing when these were most variable. This effect might just be a result of the receiver's perceptive system (Guilford & Dawkins 1991), and does not imply that some variables are more truthful or informative than others in what information they convey. Besides, females who use rapidly assessable cues will do better, as speed of choice is important in a situation of competition between females for males (Dale et al. 1992). The fact that female choice in 1995 and 1996 was best explained by two different song factors which are uncorrelated with each other, suggests that the multiple signals that can be derived from the song of this species are not redundant.

An unexpected result is the positive correlation found between the predictive value of male characteristics and the extent of female competition. It has been shown in pied flycatchers that in situations of strong female competition females make faster choices, and visit fewer males (Dale et al. 1992). This is adaptive in terms of choice strategy, as the probability of not pairing at all increases with female competition, and not breeding in a particular year will be heavily selected against in a short lived species. However, the highest predictive value of the male characteristics in this study was found in situations of high female competition. This can be interpreted if there are other territory or male quality cues which I did not measure, which are more important for a female, and which take longer to assess. The male characteristics examined in the study (song characteristics and whatever character signals arrival date) can be quickly assessed and thus become more useful when a faster choice is necessary.

The results of this study suggest that in order to understand mate choice is necessary to measure multiple male characteristics, and examine the pattern of pairing in several years in relation to other potentially important variables.

Chapter 7

EXTRA-PAIR PATERNITY AND SONG CHARACTERISTICS

Abstract

Microsatellite markers were used to determine the level of extra-pair paternity (EPP) in a Scottish population of willow warblers. The overall frequency of EPP (18% nestlings, 47% nests) is intermediate with respect to previous studies on the species. I looked at how several male variables correlated with within-pair paternity. Older birds were less likely to be cuckolded than young ones. Within-pair paternity was not related to arrival or pairing dates. Of the song repertoire characteristics, only song length was related with within-pair paternity: males with longer songs were less likely to be cuckolded. Level of genetic relatedness between the members of the pair did not influence the frequency of EPP in the nest. Survival of EPP chicks was not different from that of legitimate ones.

Introduction

The discovery that extra-pair paternity (EPP) is common has dramatically changed our understanding of mating systems in birds (Birkhead & Møller 1992). Evidence so far points to a panorama of mixed reproductive strategies, with each of the sex apparently behaving so as to gain benefits by accepting or pursuing extra-pair copulations (EPCs). Obviously, the benefit for males is increased fitness, by having more direct descendants, and there is no theoretical problem in explaining their pursuit of EPCs. The problem arises when it comes to females. How does a female benefit from multiple paternity? A first explanation from quantitative genetics simply denies the need to propose any benefits. According to this null model, if multiple copulatory behaviour is strongly selected in males, females are expected to express the same genes to some extent (Halliday & Arnold 1987). Another non-adaptive possibility is that females are forced to accept EPCs. This is clearly the case in some species, particularly in waterfowl (McKinney et al. 1983), but there is good evidence in other species that females actively seek EPCs (Houtman 1992; Neudorf et al. 1997). Two sets of ideas summarise the proposed benefits to females (Birkhead & Møller 1992):

1. Direct benefits: These include fertility insurance, paternal care of nestlings, resource acquisition, territorial defence and avoidance of retaliation by turned-down males
2. Genetic benefits: These fall into three main groups: genetic diversity, attractiveness genes or good genes, and optimal outbreeding (Rätti et al. 1995).

Research on EPP in birds has grown exponentially since the late 80s (Westneat & Webster 1994) and we now have evidence of the existence of some of these benefits. In the dunnock *Prunella modularis*, a polyandrous species, females increase their reproductive success by getting paternal care from several males, and they obtain this by mating with them (Davies 1992). It is more difficult to see what direct benefits females may obtain from multiple paternity in monogamous species, and the evidence on this is very scant. However, it has been shown that female red-winged blackbirds *Agelaius phoeniceus* have access to the territories of the extra-pair males for foraging, and they also benefit from nest defence by these males (Gray 1997). Genetic benefits in the form of good-genes have been suggested in several cases, but direct evidence for this is also scant. Kempenaers et al. (1992) have shown increased survival of extra-pair young (EPY) in the blue tit *Parus caeruleus*. However, a later research by the same authors on the same species could not find a significant effect (Kempenaers et al. 1997). A recent study in the collared flycatcher *Ficedula albicollis* (Sheldon et al. 1997) revealed a correlation between the body condition of the extra-pair young (EPY) and the size of a the white forehead patch in the EP father, which is a sexually selected character. Other studies have shown that EP males tend to have higher values in sexually selected characters which, in turn, are correlated with measures of fitness (e.g. Graves et al. 1993; Hasselquist et al. 1996; Saino et al. 1997).

The rules and outcomes of EPP games seem to be extremely varied, and explanations that fit some species are not always applicable to others, or even to the same species in different populations. For instance, the willow warbler was considered a textbook example of sexual fidelity (Gyllensten et al. 1990) until recent studies in other populations revealed significant levels of promiscuity (Bjørnstad & Lifjeld 1997; Fridolfsson et al. 1997). These population differences can be important tools, as they may allow us to take a comparative approach and so tease apart the contribution of different costs and benefits for the evolution of EP mating behaviour.

Searcy (1992) proposed that song repertoires may be used by females to choose males in the context of extra-pair copulations, specially in species where song repertoires do not seem to be used in social mate choice. One of these cases is the willow warbler, where variation in the song repertoire of which does not explain patterns of female choice in the

field (Chapter 6). In this study, I have looked at the patterns of EPP in a Scottish population of willow warblers in relation to a series of male characteristics. The objectives of the study were, firstly, to compare the population average of EPP with three earlier studies on this species and, secondly, to see if male song characteristics explain patterns of within-pair paternity.

Methods

Between 1994 and 1996, blood samples from a total of 56 clutches were collected in Kippo Wood (Fife, Scotland). Samples from the social parents at these nests could only be obtained for 34 clutches. Most males were ringed and extensively studied, so arrival, pairing and hatching dates are known. In all nests, the social father was determined by checking the identity of the male that was feeding the chicks and defending the nest with alarm calling. No cases of mate-switching occurred: all these males were owners of the territory where the female built the nest. The sampling of potential extra-pair males was not exhaustive in any year, so assignments of extra-pair paternity are quite limited.

Song was recorded from most males, and analysed following the methodology outlined in Chapter 1. Five characteristics of the song were extracted from this analysis: 1) Repertoire size; 2) Song length; 3) Song versatility; 4) Element rate; and 5) Repertoire fixation (see Chapter 1 for details).

Blood samples were taken when the chicks were 5-10 days of age. They were taken by venipuncture of the brachial vein, stored in 20 µl of 500mM EDTA and frozen in situ in a liquid nitrogen flask. The samples were stored at –70°C. DNA was extracted using a Chelex™ (Biorad) extraction method (Ellegren 1992).

Three willow warbler microsatellites were used for paternity exclusion and assignment (PHTR1, PHTR2, and PHTR3; Fridolfsson et al. 1997). The other microsatellite developed by these authors (PHTR4) could not be amplified in this population. The option of using any of the four microsatellites developed for *Phylloscopus occipitalis* (Bensch et al. 1997), was discarded, as these loci presented low polymorphism or a high incidence of null alleles. However, an additional locus from the pied flycatcher *Ficedula hypoleuca* (PTC3: Ellegren 1992) was amplified in some broods.

PCR amplifications were performed on a PTC-100 thermal cycler (MJ Research Inc.). Reaction volumes of 20 µL included 3 µL of the DNA extraction, 200 µM of each dNTP, 1.5 mM MgCl2, 10 pmol of each of the primer pair and 0.15 units of Taq polymerase (Promega). The standard Promega PCR 10x buffer was used for the reaction. The thermal treatment for PHTR1, PHTR2 and PHTR3 consisted of a first cycle at 94°C for 3 min, 56°C for 30 s, and 72°C for 40 s, followed by 30 cycles of 56°C for 30 s, 72°C for 40 s and 94°C for 40 s. A terminal extension step of 72°C for 5 min completed the program. Cycling conditions for PTC3 were identical except for an annealing temperature of 54°C. The presence of successful PCR products were checked in 3% agarose gels stained with ethidium bromide. The sizing of the fragments for all adult birds was done by electrophoresis in silver-stained 6% polyacrylamide gels for each of the loci. A 10 bp ladder (GIBCO BRL) was run every six lanes, in order to make sizing reliable. In this way, all the alleles for a given locus could be seen in relation to each other. All alleles were run in a ladder fashion in order to test the accuracy and the sizing and assess the pattern of slippage of stuttering bands, which sometimes can help in the sizing of fragments, as they appear at a consistent distance from the true allele (Fig. 7.1).

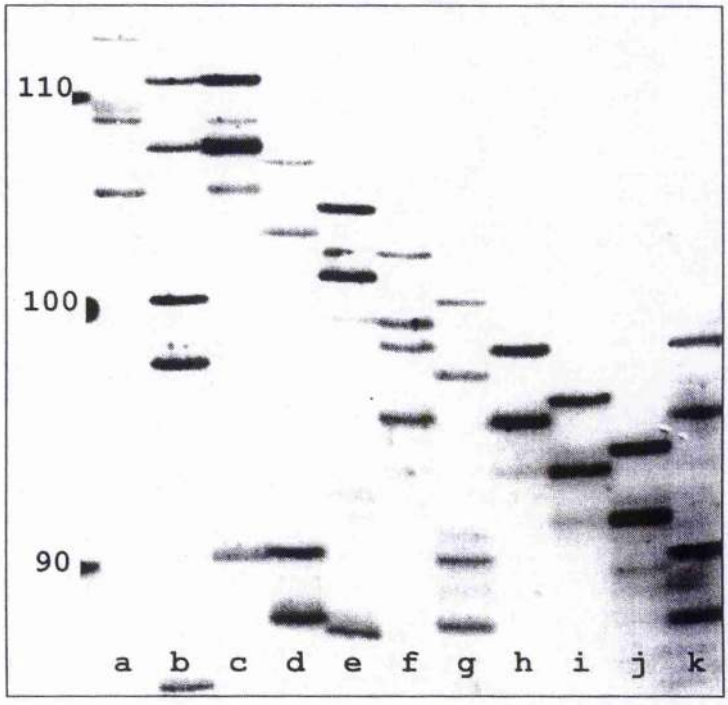


Figure 7.1. Ladder of PHTR2 alleles, with a 10 base-pair marker on the first lane from the left. In any group of several bands, the true allele is the first one seen at the top. In some cases, the two alleles are similar in size, and stutter bands overlap. Note how stutter bands appear consistently at a distance of 3 bp and 20 bp under the true allele. The lengths of the alleles (in number of base pairs) are: a) 113-109; b) 111-101; c) 111-109; d) 107-107; e) 105-103; f) 103-99; g) 101-91; h) 99-99; i) 97-97; j) 95-95; and k) 99-91.

In order to minimise errors, putative parents were run together with their nestlings. Again a ladder was run between each family in the gel. In PHTR3, sizing was more complicated due to the presence of alleles of sizes which did not match the expected pattern. All allele sizes should be odd numbered in this microsatellite. However, several even-sized alleles were found (see Fig 7.2c). I made sure that this was not an artefact by running consecutive alleles together (i.e. the presumed 144 bp allele was run with 143 and 145 at each side). Another unexpected sizing is that of PHTR1, where all allele sizes were odd numbered, contrary to the expectations based on data from the population from which this locus was described (Fridolfsson et al. 1997). These phenomena are probably due to single step replication slippage mutations in the flanking regions, which seem to be common in other microsatellites (Ji et al. 1996), and which cast doubts on the usefulness of these markers for population genetic studies. However, this does not affect their value in parentage analysis.

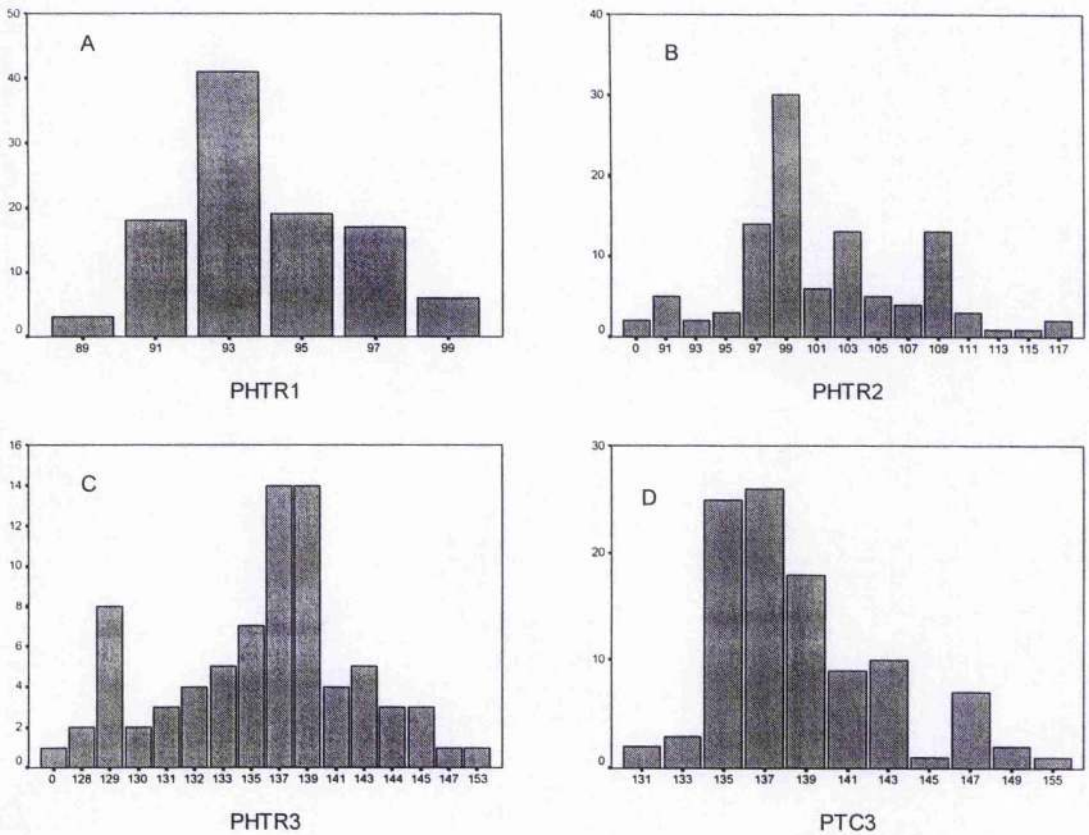


Figure 7.2. Relative frequencies of alleles at the four loci used in this study: PHTR1 (a), PHTR2 (b), PHTR3 (c) and PTC3 (d). The sample size is 52 adult birds.

Exclusion probabilities as well as Hardy-Weinberg equilibrium tests were calculated using the program CERVUS (Marshall et al. 1998). None of the loci departed from Hardy-Weinberg expectations, although the analysis of PHTR2 reveals low levels of heterozygosity and suggests the existence of a fair percentage of null alleles (Table 7.1). The

calculation of the exclusion probabilities for PHTR3 was complicated due to the fact that this locus is sex-linked in the Z chromosome, so females have only one allele. However, I was able to distinguish between male homozygotes and females by using an avian sex determination primer set (Griffiths et al. in press), and this increased the accuracy of PHTR3 for paternity testing to that of an autosomal locus.

	number of alleles	null allele frequency estimate	heterozygosity		exclusion probability
			observed	expected	
PHTR1	6	0.0406	0.692	0.758	0.535
PHTR2	14	0.1311	0.634	0.858	0.711
PHTR3	15	*	0.880	0.906	0.776
PTC3	11	0.0257	0.865	0.835	0.660

Table 7.1. Polymorphism data for the four loci used in the study. Exclusion probabilities were calculated using CERVUS (Marshall et al. 1998). In the case of PHTR3, the exclusion probability is calculated assuming that the sex of the nestling is known. The null allele estimate cannot be given for this locus, as only males can be used, and these were fewer than the minimum sample required for the analysis. Sample sizes are 52 individuals for all loci (25 males and 27 females).

In order to minimise the number of false positives, the following procedure was followed in the analysis: 1) all individuals were typed for PHTR1, PHTR2 and PHTR3; 2) if a parent-offspring mismatch was found for any given allele, an additional locus (PTC3; Ellegren 1992) was amplified for the parents and the mismatching offspring; 3) the same locus was run instead of PHTR2 in families in which there was evidence of null alleles at PHTR2; 4) a conservative rule was adopted, by which no paternity exclusion was scored if the only mismatch between putative father and offspring was a case of both being seemingly homozygotes for two different alleles each (see Fig. 7.3 for explanation).

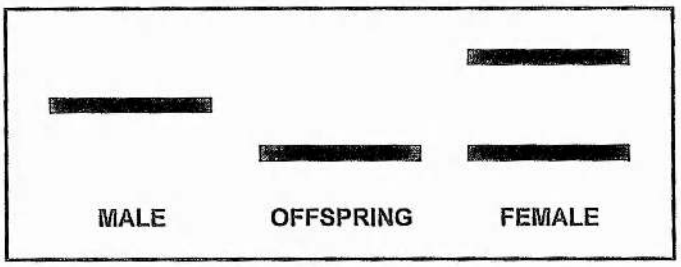


Figure 7.3. A possible case of false exclusion due to the presence of a null allele in the male genotype. If both offspring and putative male are homozygotes, then it follows that the offspring in the middle lane should be the product of an EPC. However, if the male is not a true homozygote, but has a null allele at that locus, the offspring could also share that allele and be his real offspring. Cases like this (n=4 for PHTR1, n=12 for PHTR2) were not considered mismatches unless there were additional mismatches at other loci were found.

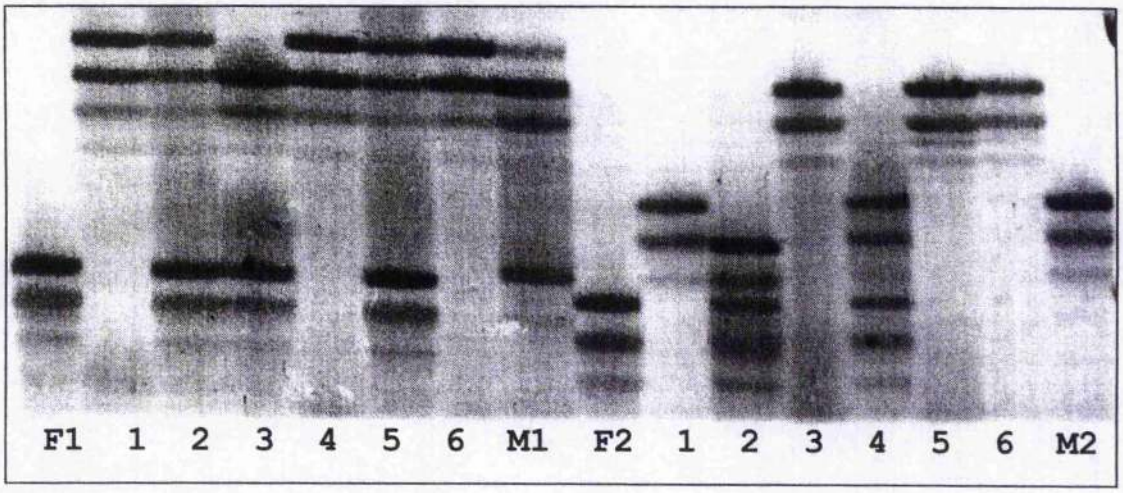


Figure 7.4. Fingerprint of the offspring of a polygynous male (M1) in 1995 in his two nests. The microsatellite shown is the sex-linked PHTR3, which shows two bands in males and one band in females. The nestlings of the primary nest (first set of 1-6) were mothered by F1, and the paternity of all of them was assigned to M1. In the secondary nest (second set of 1-6), laid by F2, an additional non-territorial male (M2) was observed feeding the chicks. In this second nest, three chicks were fathered by M1 (3,5,6), two by M2 (1,4) and another one (2) by a third male which could be identified as a neighbouring male (#4, genotype not shown). *(For a contrast between reality and fiction compare Figs 7.3 and 7.4)*

Most families were genotyped using PHTR1, PHTR2 and PHTR3; for these, the joint exclusion probability is 0.97. In some families PHTR2 was substituted by PTC3, and in this case the joint exclusion probability was 0.96. In all those chicks in which a non-parental allele was found, PTC3 was amplified in order to increase the resolution of the paternity assignment. As the population was not sampled extensively, in very few instances could the extra-pair male be identified (see Fig. 7.4 for an exception). Assignments of paternity other than by the social male were done by searching in the genotype database for males which could be consistent with the paternally inherited genotypes of the given extra-pair young (EPY). The probability of this paternity assignment being correct was calculated by using the formula $p=(2f_A-f_A^2)\ (2f_B-f_B^2)\ldots.(2f_N-f_N^2)$, where A, B.. N are the paternally inherited alleles at each locus, and f_N is the frequency of the of the N allele in the population (Jeffreys et al. 1992). Effectively, this is the probability of any other male in the population having the same genotype as the suspect male.

I used a randomisation procedure to test whether the frequency distribution and variance of EPY in the nests departed from a binomial distribution. This was done by randomly redistributing the chicks over the observed broods 100 times, and calculating the resulting percentage of EPP. I used the software RESAMPLING STATS for this purpose. The probability of a given outcome is directly calculated by dividing the number of runs in which a value greater than the observed one is found by the total number of runs.

Two measures of genetic distance between individuals were used. The genetic similarity between parents was calculated based on the sharing of alleles for the three autosomal loci, using a Jaccard similarity index: $r_{(xy)} = a/(a+b+c)$, where a is the number of alleles shared across the genotype by two individuals x and y, b is the number of alleles unique to individual x, and c the number of alleles unique to individual y. This is a standard measure of genetic relatedness, where sibs should have an average of r=0.5. A second measure based on the distance D^2 (Goldstein et al. 1995) and its derivation d^2 (Coulson et al. 1998) was developed in order to consider genetic distances between individuals at deeper branches of the pedigree. This distance should reflect phylogenetic or population differentiation, assuming a stepwise model of microsatellite evolution (Valdes et al. 1993). It was calculated as the mean squared distance in repeat units between the alleles of two individuals at a locus (i.e. four pair-wise size differences between alleles), averaged over all loci, and denoted by Δ^2, i.e.

$$\Delta^2_{(xy)} = \sum_{i=1}^{n} \frac{\left(\left|l_a - l_c\right| + \left|l_a - l_d\right| + \left|l_b - l_c\right| + \left|l_b - l_d\right| / 4 \right)^2}{n}$$

where l_a and l_b are the lengths of the two alleles of individual x at locus i, and l_c and l_d are the lengths of individual y at the same locus, and n is the total number of loci. Both measures are correlated negatively (r=-0.417, n=21, p=0.060).

Similarly, in the analysis of heterozygosity, two measures were taken. The first one is the standard heterozygosity, which is the number of loci at which an individual is heterozygous divided by the total number of loci. The second one is d^2 (Coulson et al. 1998) which reflects heterozygosity at a deeper level, and it is more a measure of the population mixing of an individual's pedigree.

Results

Frequency of extra-pair paternity

Of the total of 34 nests studied over the three years, 47% contained at least one EPY. In terms of nestlings, 18% of them were EPY. Cases of intra-specific brood parasitism (IBP), where the offspring is unrelated to both male and female, affected 8% of the nests and 2.5% of nestlings. The frequency of extra-pair paternity did nor differ between the three years (F=0.23, df=2, p=0.794). In 87% of the nests with EPY, only one extra-pair male

could be inferred from the genotypes of the EPY, in the remaining 13% a minimum of two extra-pair parents had obtained successful fertilisations.

The distribution of EPY in the nests was very different from the expected random distribution obtained through resampling (G=32.06, df=4, p<0.001; Fig.7.5A). The observed variance in frequency of EPY in a brood was greater than that expected ($p_{observed<expected}$=0.001; Fig. 7.5B). The observed distributions of EPY in nests approached a bimodal pattern, with many nests with no EPY and some with very high percentages (Fig.7.5A). Due to this skewed distribution, in the rest of the analyses, nests will be divided between having some or no EPP.

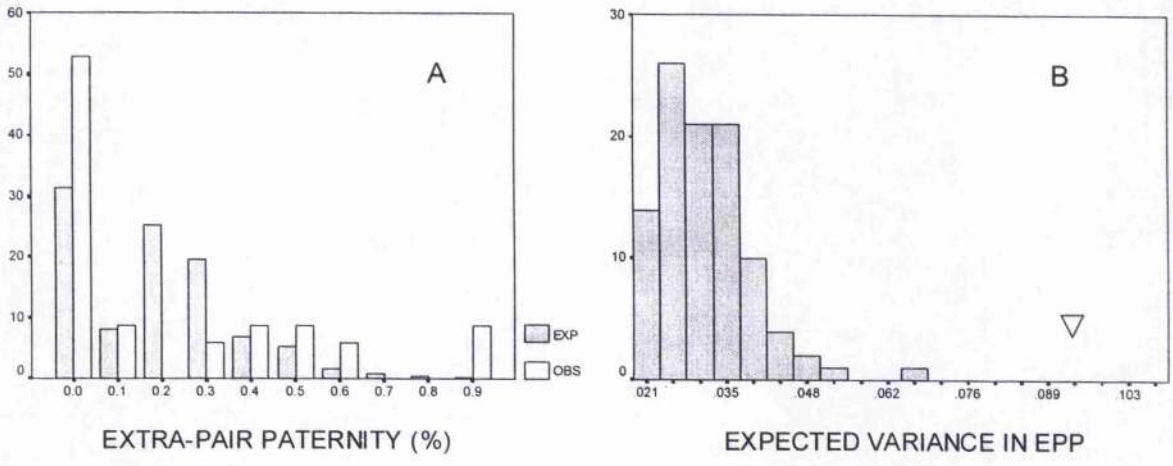


Figure 7.5. A) The observed distribution of the incidence of EPY in the populations differed significantly from the expected one after resampling the data (G=32.06, df=4, p<0.001); B) The observed variance in percentage of EPY (inverted triangle in B) was far greater than expected if EPY were distributed randomly ($p_{observed<expected}$ =0.001).

Male covariates of within-pair parentage

Variation in parentage was related to age. Males older than second-years were less likely to be cuckolded that second-years (Fisher's exact test: p=0.0238; Table 7.2).

age	cuckolded	non-cuckolded
2YR	4	9
>2YR	8	2

Table 7.2. Second-years were more likely to have EPP in their nests than older males (Fisher's exact test: p=0.0238).

Males with EPY in their nests had shorter songs (t=2.366, df=18, p=0.023; Fig. 7.6). None of the other song characteristics or dates of arrival and pairing were different between the two groups of males (all tests: -1.7<t<0.4, df=18, p>0.1).

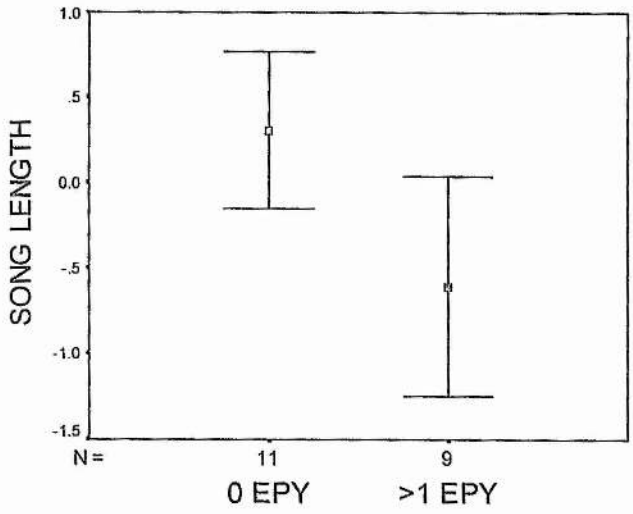


Figure 7.6. Males with EPY in their nests had shorter songs (t=2.366, df=18, p=0.023; Fig. 7.6).

In the previous analyses, data points are independent, as each males is considered only once. However, it is also interesting to know the repeatability of within-pair parentage. This can be calculated for 5 males in two consecutive years. Four of them had no EPP in their nests in either year, and the fifth one had some EPP in both years. The repeatability value is quite high (0.830). However, two of the males were paired to the same female in the two years, and this makes the data not independent.

Extra-pair paternity and male traits

There were a limited number of males which could be identified as extra-pair males (in 7 out of 16 nests with EPY). There are no obvious differences in song length between the males (Table 7.3), although it is worth noting that only in 1 out of 7 cases was the cuckolder male older that the cuckolded, and this exception was in a secondary nest. The seven males identified as cuckolders suffered no EPP in their own nests. The probability of a random sample of seven males not suffering any EPY in their own nests is 0.578 (1000 resamplings), which means that they are not different from a random subset of males. Cuckolder males were not necessarily neighbours; some of them were up to 6 territories away. Similarly distant extra-territorial forays have been observed in the field (author's unpublished observations).

cuckolded male	most likely cuckolder	probability of incorrect assignment	female status	distance in territories	cuckolder's song longer than cuckolded's?	cuckolder older than cuckolded?
85	53	0.0114	primary	6	NO	YES
79	88	0.0001	primary	5	YES	SAME
1	21	0.0052	primary	1	NO	YES
10	4	0.0047	primary	3	YES	YES
2	25	0.0001	primary	4	YES	SAME
13	4	0.0022	secondary	1	NO	YES
4	108	0.0001	secondary	4	YES	NO

Table 7.3. List of successful extra-pair parentage assignations, with probabilities of incorrect assignments.

Genetic similarity between pair members and EPP

I tested the possibility that the degree of EPP in a clutch was related to the degree of relatedness (r) or genetic distance (Δ^2) between the members of a pair. Firstly, I tested whether there was assortative mating based on these two measures of genetic similarity. The values observed in the population for r (mean=0.21, sd=0.14) and Δ^2 (mean=9.00, sd=5.21) fall within the confidence interval of the mean of the expected distributions obtained through resampling. Thus, there was no assortative mating with respect to any of the two measures of genetic similarity (r: $p_{observed>expected}=0.670$; Δ^2: $p_{observed>expected}=0.396$). Figure 7.8 shows the expected distribution for the values of Δ^2 after pairing the genotyped individuals randomly 1000 times.

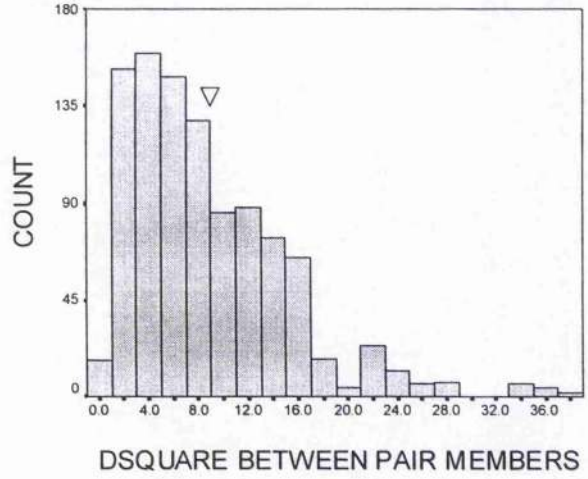


Figure 7.8. After randomly pairing the adult birds in the population 1000 times it is possible to calculate the Δ^2 which the two members of a pair would have in the absence of assortative mating based on that variable. The observed value (marked by an inverted triangle in the figure) does not differ from the confidence interval of the mean of the expected distribution ($p_{observed>expected}=0.396$).

The level of within-pair paternity in a clutch is not related to the genetic similarity between the members of the social pair as measured by either the degree of allele sharing or the genetic distance between the two parents (r: $t=0.408$, $df=19$, $p>0.05$; Δ^2: $t=1.056$, $df=19$, $p>0.05$).

Sex of EPY

It has been predicted that if the benefits behind EPP are genetic, females should bias the sex ratio of their EPY in favour of males (Sheldon & Ellegren 1996). The same bias can be expected for chicks resulting from intra-specific brood parasitism considering that nestling males are heavier (Chapter 8), and probably more costly to rear. If we compare the sex distributions of all the non-legitimate (EPY and IBP) with the legitimate young, we see that non-legitimate young are not more likely to be male than female ($p=0.223$, Fisher's exact test). Similarly, a matched pair comparison between the sex-ratios of EPY and legitimate young in those nests with at least one extra-pair chick revealed no differences (Wilcoxon matched pairs test, logit transformed data, $Z= -1.12$, $n=13$, $p = 0.270$). However, the trend in sex-ratio is in the expected direction (see distribution of sexes between the three classes in Table 7.4)

	Males	Females	Total
Legitimate	75	84	159
EPY	20	16	36
IBP (egg dumping)	4	1	5
Total	99	101	200

Table 7.4. Distribution of sexes in the whole sample, divided into legitimate, extra-pair-male young and offspring resulted off intra-specific brood parasitism.

Offspring survival and EPP

Unfortunately, data on the body condition of nestlings were not taken, so I cannot compare whether legitimate and EPY differed in this fitness component (Sheldon et al. 1997). However, I can test whether there are differences in the probability of recruitment into the population from those nests that were not predated before the young fledged, although the small sample size makes this test not powerful enough. As males are more philopatric than females, differences in recruitment could be sex specific. However, EPY are not more likely to recruit than legitimate young regardless of whether each sex is considered separately or the two are pooled together (Fisher's exact tests; all $p>0.05$; Table 7.5).

	Males		Females		Total	
	Recruits	No	Recruits	No	Recruits	No
Legitimate young	13	54	7	74	20	128
Extra-pair young	3	17	1	14	4	31

Table 7.5. Figures of recruiting and non-recruiting nestlings divided by sex and paternity. There are no differences in the probability of surviving (15% overall for the two sexes).

If we look at individual nests, there are no differences between cuckolded and non-cuckolded nests in how many recruits they produce (Mann-Whitney U=96.0, n=30, p=0.56). Cuckolded nests were no more likely than non-cuckolded nests to have at least one nestling recruiting (Fisher's exact test p=0.13), though there is a trend in that direction (Table 7.6). However, in those cuckolded nests with at least one recruit, the recruiting young was not more likely to be extra-pair than legitimate (4 extra-pair, 5 legitimate; binomial test, p=1.0).

	At least one recruit	No recruits
Cuckolded	10	3
Non-cuckolded	9	7

Table 7.6. The probability of having at least one nestling recruiting is not different for cuckolded and non-cuckolded nests (Fisher's exact test p=0.13).

Genetic variation and EPP

It has been shown recently for two mammal species that genetic diversity as measured by mean d^2 and heterozygosity is a good predictor of an important component of fitness (Coltman et al. 1988; Coulson et al. 1998). Following this argument, it can be predicted that if d^2 is also correlated with fitness in this species, females paired to males with low values of d^2 or low heterozygosity should have more EPP than those paired to males with high genetic variation. The results do not support this hypothesis: neither heterozygosity or d^2 of the social male were different between nests with and without EPY (both $t<0.719$, $n=19$, $p>0.5$).

I tested whether male chicks with high d^2 values were more likely to recruit into the population. A logistic regression on the whole sample revealed no predictive value of d^2 ($\chi^2=0.23$, df=1, p=0.62). If we consider only those nests with some recruiting offspring, a

matched pairs test between the average d^2 of recruiting and non recruiting birds also fails to find any significant difference ($t=1.57$, $df=17$, $p=0.135$).

Polygyny and EPP

In the total sample three males became polygynous, by having attracted a second female sometime during the laying period of the first one. One male was polygynous in two consecutive years. The sample size is very small but it is worth comparing the levels of EPP in primary and secondary nests (Table 7.7). We can see that in three out of four cases, primary females had no EPP, and secondary ones did. The pattern is inverted in the fourth case. In the secondary nest of 13 in 1995, a non-territorial second male shared paternity (0.33) in the secondary nest, and was observed feeding the young.

Male and year	within-pair parentage in the nest of the primary female	within-pair parentage in the nest of the secondary female
4 (1996)	1	0.4
13 (1995)	1	0.5
13 (1996)	1	0.6
52 (1996)	0.6	1

Table 7.7. Cases of polygyny in the sample and pattern of within-pair parentage.

Intraspecific brood parasitism

The low degree of intraspecific brood parasitism (IBP) allows very little scope for further investigation. In only three nests could IBP be detected, and each of these was found in a different year. However, it is interesting to note some peculiar facts about these nests. Firstly, two of them were "socially" fathered by the same male (#2) in two consecutive years and in the same territory. These two nests were built by two different females, but almost in the same place (within two meters of each other). The other female that was subjected to egg dumping bred in the same territory with the male that took it over from male #2 in 1996, and her nest was built within five meters of the place where her predecessor had built hers. Given that all broods considered in this study were sampled in an area of more than 9 Ha., this local clustering of IBP is unlikely to be a coincidence. One possibility is that a single female breeding in the area had parasitised all three nests in the three years. However, the genotypes of the five IBP chicks are not compatible with the

hypothesis of a single female having laid them all. Another possible explanation could be that the habitat where these nests were built makes nests easy to find for females. Although the habitat around this particular territory does not differ strikingly with other areas of the wood, the vegetation is more open than average, and there is a wide track and a road nearby, which may increase the chances of females being observed while nest building.

The three nests with IBP also contained extra-pair young, at a higher percentage than average (Mann Whitney's $U=90$, $df=34$, $p<0.038$).

Another interesting fact is that one of the nests with IBP was a very late replacement clutch, laid after the first clutch had been preyed upon while containing young almost ready to fledge. At the time when this replacement clutch was being laid all the rest of the nests in the population had young already. This could indicate that females might be able to produce eggs in response to a sudden opportunity.

Discussion

Frequency of EPP: populational patterns

The first study to look at extra-pair paternity in the willow warbler was conducted in central Sweden (Gyllensten et al. 1990) and did not find a single EPY in a sample of 19 families. However, recent studies in other populations have given a different picture, with frequencies of 28% EPY in the Swedish island of Gotland (Fridolfsson et al. 1997), and 33% EPY in South Norway (Bjørnstad & Lifjeld 1997). My results in this Scottish population yield an intermediate figure of 18% EPY. What is the cause of these population differences?

A study on eastern bluebirds *Sialia sialis* (Gowaty & Bridges 1991) demonstrated experimentally that high breeding densities increased the levels of EPY and IBP. This study was the first demonstration that the adoption of alternative strategies such as EPCs were dependent on the occurrence of opportunities for adopting them. A recent review (Westneat & Sherman 1997) confirmed this pattern: in a within-species analysis, the frequency of EPY in a given population correlated with breeding density. However, in the case of the willow warbler, even if a high breeding density could be a necessary condition for EPP to occur, it still does not explain the pattern on its own (see Table 7.8).

population	% EPY	density (pairs/km^2)	reference
Mainland Sweden	0%	58	(Gyllensten et al. 1990)
Gotland	28%	80	(Fridolfsson et al. 1997)
South Norway	33%	140	(Bjørnstad & Lifjeld 1997)
Kippo (East Scotland)	18%	190	this study

Table 7.8. Breeding densities and levels of EPY in the four studies conducted so far in willow warblers.

Another factor which has been shown to have an influence in the frequency of EPY in different populations is breeding synchrony. However, there is controversy as to the direction of this effect. It has been argued that high breeding synchrony reduces the opportunity for EPCs, as males will have to guard their mates at the time that others are fertile (Birkhead & Biggins 1987). Stutchbury & Morton (1995) proposed on the contrary that breeding synchrony would make it easier for males to find fertile females, and that it would also enable females to estimate male quality and thus exercise their mating preferences. In the blue tit, it has been found that greater synchrony does indeed result in higher rates of EPY (Kempenaers 1997). However, loss of paternity in that study was not related to whether the male engaged in EPCs during or after the fertile period, which suggests that the female choice is only slightly constrained by mate guarding.

Fridolfsson et al. (1997) suggested that in the willow warbler, open habitats with good visibility may facilitate EPCs by allowing females to choose mates. However, female willow warblers probably do not use sight but song to locate males. Density of vegetation is bound to have an influence in EPP, although it is difficult to predict in which direction. Openness of vegetation will ease mate guarding, particularly in populations of low breeding density. Close habitats with low visibility may probably allow females to escape mate guarding more easily. Similarly, males seeking EPCs will be less likely to be detected by mate-guarding males when on extra-territorial forays. Although open habitats will make mate guarding more easy, they will also allow males seeking EPCs to find females.

The most likely explanation for a the variation in the levels of EPP in different populations is based on differences in density, habitat and breeding synchrony. These variables affect directly the effectiveness of mate guarding (Birkhead & Møller 1992) and the choice of EP males that a female can achieve. However, important variations between years have been found which do not seem to respond to changes in habitat variables (Graves et al. 1993;

Dunn et al. 1994; Barber et al. 1996). Petrie & Kempenaers (1998), propose several possibilities based on good-genes hypotheses to explain variation between populations in patterns of EPP. They suggest that variation in EPP may correlate with the level of genetic diversity in a population. One prediction is that EPP should be less frequent on islands and in populations which have gone thorough a bottleneck. The data from the Gotland population (Table 7.2; Fridolfsson et al. 1997) does not support this prediction. We do not know however what levels of genetic variation exist in the different populations. Another problem is to decide what we mean by genetic variation, and at that level this is measured. It seems more parsimonious to assume that variation in EPP in different populations of the same species is mostly influenced by differences in how easy it is for females to look for EP partners, and for males to stop them from doing so. However, if females accrue genetic benefits through EPP, populational differences in the variance of male quality could influence the level of EPP in the population (Petrie & Kempenaers 1998).

Parentage: individual

If EPP were only influenced by population patterns, we would expect EPY to be randomly distributed in the population. This was not the case, as there are far too many nests with either no EPY or a very high frequency of EPY. This frequency distribution is typical of EPP (e.g. Lijfeld et al. 1993; Burley et al. 1996), and it is not specifically predicted by any particular hypothesis of female benefits. Although it has been suggested that, if genetic variability benefits were important, all nests should show the same level of EPP (Petrie & Kempenaers 1998), this argument ignores the fact that mate guarding abilities might differ between males and that female choice of EP partners can be constrained. This argument will be considered later in the discussion.

Old males were less likely than young ones to suffer loss of paternity. Higher confidences of within-pair parentage with increasing age have been found in some species (e.g. purple martins *Progne subis* (Wagner et al. 1996) or American redstarts *Setophaga ruticilla* (Perreault et al. 1997)). However, in many species, old males suffer similar levels of EPP in their nests to young ones (e.g. yellow warbler *Dendroica petechia:* Yezerinac & Weatherhead 1997; blue tit: Kempenaers et al. 1997; *Ficedula hypoleuca:* Rätti et al. 1995; red-winged blackbirds: Weatherhead & Boag 1995), or are even more likely to be cuckolded (indigo bunting *Passerina cyanea:* Westneat 1990). In the house sparrow *Passer domesticus*, and in the yellowhammer *Emberiza citrinella*, although old males are more likely to obtain EPP outside their nests, they are as likely as young birds to suffer it in their own nests (Wetton et al. 1995; Sundberg & Dixon 1996). It is expected that if females are looking for good-genes,

and if there is a additive genetic variance in survival, females would benefit by choosing long-lived males as father of their offspring (Trivers 1972; Manning 1985). This is consistent with the results from the paternity assignments of this study: very rarely an EP male was younger than the social male. This could either imply that females are accepting/seeking EPCs from older males than their social one, or that older males are better at mate guarding.

Neither arrival date or pairing date correlated with levels of EPP. This was unexpected because arriving date is strongly correlated with pairing speed, and we would expect high quality males to be chosen first by females and to have higher within-pair paternity than late ones. This can be also predicted by the fact that more males will be free of mate guarding chores late in the season and thus there should be a higher opportunity for EPCs. However, males were as likely to be cuckolded irrespective of when they arrived or paired.

Among the male song characteristics, only song length was related to within-pair paternity: females paired to males with long songs were less likely to have EPY in their nests than those paired to males with short songs. This was not a consequence of young males having shorter songs, as average song length does not change with age (Chapter 3). In the same population, song length did not explain the pattern of pairing. This lack of concordance suggests that benefits derived from social mate choice are different to those accrued by EP mate choice. If song length is an indication of male quality, the results would imply that females paired to low quality males are seeking/accepting EPCs from other males. As song length correlates negatively with survival in second-years (Chapter 5), this can be interpreted as males experiencing the costs of a sexually-selected handicap (Zahavi 1975; Maynard Smith 1987). However, if females were choosing EP males on the basis of genetic benefits, we would expect that those characteristics that predict offspring viability would be related to the level of EPP. In this study I have shown that the only male characteristic that predicts offspring viability to some extent is repertoire size (Chapter 5). However, repertoire size was not related to extra-pair paternity. This cast some doubts on the previous suggestion that genetic benefits could be behind the higher levels of EPP in the nests of males with short songs.

Searcy (1992) suggested that song repertoires might be of selective advantage to males in the context of EPP. This argument was based on the disparity between the strong female preference for song repertoire found in lab experiments, and the lack of relevance of repertoire size as a predictor of pairing in many field studies. The only study that has looked at this suggestion so far is that of Hasselquist et al. (1996), who found that in the

great reed warbler *Acrocephalus arundinaceus*, cuckolders had a larger repertoire than the male they cuckolded. Repertoire size in that species is a good predictor of pairing success and viability, and thus it seems that females are obtaining genetic benefits from choosing EP males with large repertoires. However, this was not a test of Searcy's prediction, as the great reed warbler is one of the species in which mating success is explained by repertoire size (Catchpole 1986). The lack of significance of repertoire size at explaining mate choice and EPP patterns in the willow warbler suggest that a functional explanation for the maintenance of song repertoire in this species should address the effects of this trait on intrasexual selection (Chapter 9)

All the patterns found so far (non random distribution of EPP, higher paternity with age and shorter song length of cuckolded males) are only correlational, and this does not demonstrate that they are a consequence of active female choice. The same results could come about if, for instance, age or song length covary with intensity of mate guarding. It is known that male size influences the success of mate guarding in the purple martin (Wagner et al. 1996), and that experimental male detention results in increased levels of EPP in western bluebirds (Dickinson 1997). However, it is also known that intensity of mate guarding does not always correlate negatively with EPCs. For instance, in the blue tit (Kempenaers et al. 1995), the percentage of EPY in a nest was not related to the intensity of mate guarding performed by the male. Dunn et al. (1994) have shown that, in the tree swallow *Tachycineta bicolor*, females seem to be in control of the extra-pair fertilisations regardless of the possible ecological constraints of density and synchrony. It is difficult to know what the situation is in the willow warbler. Although the density of vegetation suggests that females should be able to avoid mate guarding and engage in EPCs, we do not know what costs this behaviour may have (e.g. smaller feeding effort by social partners).

Repeatability of within-pair paternity in different years was very high, although some males paired with the same female and data are thus not independent. High repeatabilities in within-pair paternity have been found in most species where data for several years or seasons were available (Weatherhead & Boag 1995; Kempenaers et al. 1997; Møller & Tegelström 1997).

Relatedness and EPP

One of the possible benefits behind EPP is that females paired to too closely or distantly related males may obtain optimal levels of outbreeding by choosing a different father to the social male (Rätti et al. 1995). The resampling analysis shows that neither Δ^2 nor the degree

of allele sharing between the members of the pair differed from that expected through random mating. The frequency of EPY in the clutch was also unrelated to the degree of allele sharing or Δ^2 between the two members of the pair. I used a small number of loci to estimate relatedness, and this certainly renders this analysis less sensitive to the lower limit of relatedness (i.e. the outbreeding end of the continuum), so the results are not to be taken as definitive.

The only study which has previously looked at that possibility (Rätti et al. 1995) found that pairs of pied flycatchers that had very low coefficients of relatedness (estimated through sharing of bands obtained by DNA multilocus fingerprinting) were more likely to have EPY than those that had average relatedness. This was interpreted to mean that females were avoiding extreme outbreeding, and the authors presented some data which suggests hatching costs of both inbreeding and outbreeding. Although increased hatching failure associated with high levels of inbreeding has been previously documented (Bensch et al. 1994), negative effects of outbreeding have not yet been found in any natural population of birds.

Sex of EPY

Although there was a tendency for EPY and IBP chicks to be male, the difference was not statistically significant. The two previous studies that have addressed this issue have found mixed results. In the blue tit, although the overall sex ratio was not different from parity for both legitimate and EPY, a matched-pair analysis comparing the sex-ratios for individual nests revealed that EPY were more likely to be male (Kempenaers et al. 1997). The other study was on the collared flycatcher *Ficedula albicollis* (Sheldon & Ellegren 1996) and found no differences in sex-ratio related to paternity.

Theory predicts that if the benefits accrued to females through EPP are genetic, it would pay females to produce male EPY (Sheldon & Ellegren 1996). If males are more costly to rear, a male biased sex-ratio is also expected in chicks result of egg dumping. Male chicks are bigger in the willow warbler (Chapter 8), and probably more costly to produce. Furthermore, this size difference probably makes males better at competing with other nestlings than females. Female birds have been shown to be able to produce adaptive sex ratios (e.g. Ellegren et al. 1996; Komdeur et al. 1997), but the mechanism of control may not be precise enough as to bias the sex of chicks of different fathers. However, the trend found in this study goes in the predicted direction, and a bigger sample size might reveal significant differences.

Offspring survival

The only study that has shown increased survival of EPY was done in Belgium in blue tits (Kempenaers et al. 1992). However, the same authors working in the same species could not replicate the results with an increased sample size (Kempenaers et al. 1997). In the present study, EPY were as likely as legitimate chicks to survive. There was a trend for nests with EPY to be more likely to have at least one surviving chick, although this was not significant. The fact that those surviving chicks were as likely to be EPY as legitimate suggests that the effect of good-genes was not large, or that the choice of EP males by females was restricted. As it was mentioned before, male repertoire size is the best predictor of offspring viability, and this variable was not related to EPP.

Genetic variation and within-pair paternity

Genetic diversity of the social male as measured by heterozygosity and d^2 did not correlate with within-pair paternity. Chicks with higher levels of genetic diversity were not more likely to survive than those with low levels. A correlation between genetic diversity as measured by d^2 and fitness have been reported in two species of mammal (Coltman et al. 1988; Coulson et al. 1998), but no study has looked at this relationship in any bird population. It is also not very clear what d^2 means, although it seems to measure genetic distance at an almost population or phylogenetic scale. More studies are needed to see if similar patterns to those found in mammals apply to bird populations. This should provide a promising interface with studies of dispersal and philopatry.

Polygyny and EPP

Most secondary nests of polygynous males had EPY. In one of these nests, a non-territorial male was observed feeding the young. Although females can raise a whole clutch on their own (Bjørnstad & Lifjeld 1996; personal observations), paternal help promotes early fledging in this species (Bjørnstad & Lifjeld 1996), but secondary females do not get paternal care before the chicks of the primary nest fledge (personal observations). The trend for secondary females to have higher levels of EPP may reflect the attempt of these females to get extra paternal care. If the good-genes benefits derived from mating with polygynous males would outweigh direct ones, we would expect these females not to have any EPY in their nests. This prediction is met in some studies (e.g. in blue tits: Kempenaers et al. 1995, and Savannah sparrows *Passerculus sandwichensis*: Freeman-Gallant 1997), but not in many others (e.g. tree swallows: Dunn & Robertson 1993; red-winged blackbirds: Gibbs et al. 1990; or house wrens: *Troglodytes aedon*: Soukup & Thompson 1997). However, even if the variance in reproductive success of polygynous males is reduced though EPP in their

nests, it can still likely to be greater than that of monogamous males, specially if extra-pair success is taken into account (Gibbs et al. 1990; Yezerinak et al. 1995).

As not all nests and potential fathers were surveyed, it is not possible to calculate the reproductive success of the males in this population. However, polygynous males fledged fewer nestlings that than would be expected from counting the number of chicks in their nests. This cost might be balanced out if polygynous males are also better at achieving extra-pair fertilisations.

Egg dumping (IBP)

Nests which suffered high IBP also had high levels of EPY. This result can be interpreted in two ways. Males stay close to the female while she is building the nest (personal observations; Arvidsson 1992) and produce alarm calls if the observer is seen (Chapter 2), which normally result in the female stopping temporarily nest building (personal observations). Parasitic females probably locate nests by following females while they build the nest, and males might also defend their female against other females espionage attempts. If the ability of males to mate guard covaries with the ability to protect nests against other females' parasitism, we would expect to find a correlation between IBP and EPY. However, the fact that the three parasitised nests were located in the same territory in different years suggests that nest location is an important factor when it comes to IBP. If this is so, the costs of suffering IBP would select for females to build nests in places where they cannot be observed by other females. The requirements of nesting places might not be the same under predation and IBP selection pressures. For instance, if breeding density is correlated with IBP, this might be important in limiting the extent of polygyny in a population. This adds a new level of complexity to the question of what territory quality might mean (Chapter 6).

Chapter 8

SEX-RATIO VARIATION IN BROODS

Abstract

The sex of 405 willow warbler hatchlings was determined by PCR amplification of an area of a CHD gene present in the sex chromosomes. The overall sex-ratio was not different from 50:50, and the distribution of chicks of both sexes in the broods did not differ from the expected binomial model. However, there was a negative correlation between male song repertoire size and the proportion of males in a brood. Furthermore, individual females decreased the proportion of males in their brood when breeding with the same male for a second year. This is contrary to predictions based on the heritability of attractive characters. An explanation based on Local Resource Competition is tentatively suggested.

Introduction

Fisher (1930) suggested that stabilising selection was responsible for the pervasive existence of 50:50 sex ratios. He argued that if a given population was not composed of a balanced sex ratio, a mutant which would produce an excess of the rare sex would leave more descendants, and thus quickly drive the population to the other extreme. This process would eventually lead the population towards an optimum 50:50. However, he also hypothesised that when the two sexes are not equally costly to produce, the sex ratio should be skewed towards the cheapest sex, in direct proportion to the difference in cost between producing the two sexes.

This Fisherian null hypothesis has been successfully challenged in several occasions, notably because it is based on assumptions which do not always hold true. Populations are seldom infinite and sympatric, or composed of individuals of similar fitness. Besides, if a facultative manipulation of sex ratios is possible, we face a different game altogether, where individuals can change sex ratios to maximise their fitness depending on particular circumstances (Charnov 1982). In the case of birds, females are the heterogametic sex (ZW), so biases in sex ratio are expected to reflect the female's interests, given that males (ZZ) can only provide a single type of sexual chromosome.

The evidence for adaptive variation in sex ratios can be divided into several categories depending on the evolutionary pressure responsible for the adaptation:

1. If the mating or dispersal patterns of the offspring are such that close kin are likely to compete against each other for mating (Local Mate Competition) or resources (Local Resource Competition), unequal sex ratios can evolve through kin selection. Most examples of this kind are found in insects, particularly in parasitoids (Hamilton 1967), but the argument has also been used to explain the skewed sex ratios found in some mammals, like bush-babies *Galago spp.* (Clark 1978) and roe deer *C. capreolus* (Hewison & Gaillard 1996).
2. If parental condition affects male and female offspring fitness to a different degree, it is expected that parents in good condition should produce more offspring of the costlier sex (Trivers & Willard 1973). This often leads to seasonal variations in sex ratio, as parents reach their reproductive condition at different times. In red deer *Cervus elaphus* for instance, sons are more costly to produce and, accordingly, females in better condition give birth to more sons (Clutton-Brock 1991). The pattern is reversed in the case of rhesus macaques *Macaca mulatta*, where subordinate females produce an excess of sons, which are less costly to produce in that species (Gomendio et al. 1990).
3. Seasonal effects are not always due to variation in parental condition and purely environmental changes may influence variation in sex ratios. For instance tawny owls *Strix aluco* adjust the sex ratio of their broods in response to the availability of voles (Appleby et al. 1997). Seasonal trends in sex ratios are also very common in Falconiformes, but not always in the same direction. In this case it seems that the most important constraint is maturation time, which is different for the two sexes, and also varies between species. It has been shown recently that the sign of the bias in sex ratio seems to depend on which gender has its fitness more negatively affected by date of hatching in the different species (Daan et al. 1996).
4. Skewed sex ratios are often found in species which show cooperative breeding and in which the roles of helping vs. dispersal are sex determined (Gowaty & Lennartz 1985; Heinsohn et al. 1997). The most striking example is that of the Seychelles warbler *Acrocephalus sechellensis*. In this species females stay on the parents' territory and help rear subsequent clutches, whereas males disperse. Komdeur *et al.* (1997) have shown that sex-ratios in good quality territories are female biased, whereas those in poor quality areas are male biased. This strategy is adaptive, as bad quality territories cannot support helpers. These authors went further to show the same pattern experimentally, by translocating birds between islands.

5. Another situation where sex ratio adjustment is adaptive is when offspring inherit characters determining their attractiveness. Burley (1981,1986) has shown that sex ratios in zebra finches *Taeniopygia guttata* are biased towards the sex of the parent which has been artificially made more attractive with colour rings. In the same line, female blue tits *Parus caeruleus* paired to males which survive to the next year have male-biased clutches (Svensson & Nilsson 1996). In collared flycatchers *Ficedula albicollis*, where the white forehead patch of the male is a sexually selected trait, the proportion of males in a given clutch is directly related to the size of this patch (Ellegren et al. 1996). As variance in reproductive success is greater for males than for females, this is interpreted in the sense that females obtain genetic benefits for their sons, either in the way of good genes effects or in Fisherian attraction.
6. There are other variables which have been found to correlate with sex ratio in some bird species and whose function is not altogether clear. Amongst them we find: hatching date: great tits *Parus major* produce male biased clutches late in the season (Lessells et al. 1996); female age: older mothers produce more sons in red-winged blackbirds *Agelaius phoeniceus* (Blank & Nolan 1983); male arrival date: late arriving males have male biased clutches in the great reed warbler *Acrocephalus arundinaceus* (Westerdahl et al. 1997).

The mechanisms responsible for skewed sex ratios are varied, and range from differential implantation in mammals (Clutton-Brock 1991) to meiotic drive in some insects (Wilkinson et al. 1998). In birds, sex is chromosomally determined, and the development of the egg is external. This explains the traditional neglect of research into adaptive sex ratios in birds. Until now, no mechanism by which a manipulation of sex ratio could come about in birds has been fully described, although several possibilities at different stages of reproduction have been suggested (Oddie 1998). The manipulation could happen either at the time of chromosome segregation, during follicle differentiation or after ovulation. The second possibility is suggested by the fact that not all ova develop into follicles (Wood-Gush & Gilbert 1970), and by the potentially different developmental rates that Z and W follicles may have (Krackow 1995). Reabsortion after fecundation is another possibility in species where the lag between successive eggs is longer than a day (Oddie 1998). However, avian reproductive biology is poorly understood and so far all these mechanisms are nothing but speculations.

The main problem that research on bird sex ratios has encountered is the difficulty of measuring primary sex ratios, given that direct sex identification of hatchlings is rarely possible. Fledglings are typically sexed instead, but in this case differential mortality in the

nest can be mistaken for variation in primary sex ratio. However, the recent development of molecular markers which can be amplified through PCR, and which require minute amounts of DNA has made sex identification of hatchlings feasible and quick (Griffiths et al. 1996; Griffiths et al. in press). The aim of this study was to investigate whether hatchling sex ratios in the willow warbler *Phylloscopus trochilus* depart from Fisherian ratios and to analyse any skew found in relation to a set of male variables.

Methods

Between 1994 and 1997 blood samples from a total of 69 clutches were collected in Kippo Wood (Fife, Scotland). Most adults were ringed and extensively studied, so biometrical data as well as arrival, pairing and hatching dates are known. Song was recorded from most males, and analysed following the methodology outlined in Chapter 1. Five factors of song complexity and performance were extracted from this analysis: 1) Repertoire size; 2) Song length; 3) Song versatility; 4) Repertoire diversity; and 5) Repertoire fixation (see Chapter 1 for details).

Blood samples were taken when the chicks were between 5 and 10 days of age. This sample is representative of the primary sex ratio as in a total sample of 75 broods studied in this population in the same period, no natural nestling mortality was detected. The incidence of unfertile eggs is small (6.5% of broods contained one unhatched egg). Blood samples were taken by venipuncture of the brachial vein, stored in lysis buffer (500mM EDTA) and frozen in situ in a liquid nitrogen flask. The samples were kept at –70°C in the lab. DNA was extracted using a Chelex extraction method (Ellegren 1992). PCR amplifications were performed on a PTC-100 thermal cycler (MJ Research Inc.). Reaction volumes of 20 µL included 3 µL of the DNA extraction, 200 µM of each dNTP, 1.5 mM MgCl2, 20 pmol of each of the primers P8.3221 (5'-CTCCCAAGGATGAGRAAYTG-3') and P2 (5'-TCTGCATCGCTAAATCCTTT-3') (Griffiths et al. in press), and 0.15 units of Taq polymerase (Promega). The standard Promega PCR 10x buffer was used for the reaction. The thermal treatment consisted of an initial denaturing step at 94°C for 1 min 30s, followed by 30 cycles of 54°C for 45 s, 72°C for 45 s and 94°C for 30 s. A terminal cycle of 54°C for 1 min and an extension of 72°C for 5 min completed the program (Griffiths et al. in press). The PCR products were visualised in 3% agarose gels stained with ethidium bromide. In each gel run one adult bird of each sex was included. Adult birds were sexed in the field by differences in wing length and presence of a brood patch (Svensson 1984). The reliability of the molecular technique was assessed by looking at the band pattern of 32 males and 28 females sexed on these criteria.

The fragment amplified from the Z chromosome contains a restriction site which can be cleaved by *HaeIII* (Griffiths et al. 1996). I tested this pattern by digesting PCR products from a male and a female. 5 units of *HaeIII* were added to a 8 μL of PCR product in 1x Gibco restriction buffer III and 50 ng μ^{-1} of BSA in a total volume of 10 μL.

I used a randomisation procedure to test whether the frequency distribution and variance of sex ratios corresponded to that expected from a binomial distribution. This was done by randomly redistributing the chicks over the observed broods 100 times, and calculating the resulting sex ratio. The probability of a given outcome is directly calculated by dividing the number of runs in which a value greater than the observed one is found by the total number of runs. Analysis of proportions such as sex ratio is best done by using a logit transformation (Sokal & Rohlf 1995), which makes the relations between proportions and causative variables linear. Analysis of sex ratio variation in relation to a range of possible variables was performed using a generalised linear model with binomial errors and a logit link (Genstat 5 Committee 1993). The statistical significance of each of the possible correlates was tested by looking at the change in deviance from the full model when a given variable was removed from it. The significance of the change in deviance statistic has approximately a χ^2 distribution (Genstat 5 Committee 1993).

Results

All adult birds (32 males, 28 females) were classified correctly by the pattern of bands resulting from PCR amplification. Males presented a single band at around 340 bp, and females showed an additional band to that one at approximately 380 bp. Digestion by *HaeIII* resulted in the fragmentation of the Z band into fragments of 280 bp and 60 bp approximately, whereas the W band remained uncut (Fig. 8.1) as expected (Griffiths et al. 1996). The pattern of single vs double bands in the gel following amplification was thus considered sufficient for sexing of the chicks.

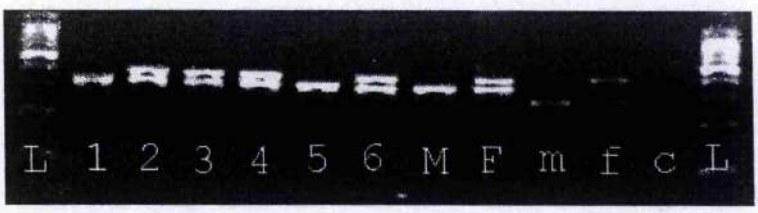


Figure 8.1. Agarose gel stained with ethidium bromide showing the amplification products of a complete family (1 to 6), with the parents (M, F). Males show one single band and females two. The last two lanes (m, f) show the *Hae*III digested male and female PCR products, revealing the presence of a restriction site in the male fragment. The ladder (L) is a 100 bp ladder (Promega), the lowest band is 100 bp.

Of the 405 chicks sexed in total, 212 (52%) were male. This is not different from the expected 50:50 ratio (G=0.89, df=1, p>0.05). The population sex ratios did not depart from the 50:50 ratio in any given year (G>2.6, df=1, p>0.05; all years; Table 8.1).

	# Males	# Females	Total
1994	40 (57%)	30 (43%)	70
1995	59 (48%)	64 (52 %)	123
1996	79 (57%)	60 (43 %)	139
1997	34 (47%)	39 (53 %)	73
All years	212 (52%)	193 (48 %)	405

Table 8.1. Numbers of male and female willow warblers sexed from the sampled nests in Kippo Wood (1994-1997). These are primary sex ratios, there being no nestling mortality.

The frequency distribution of the sex ratio in the broods did not differ from that expected from a binomial distribution (Fig 8.2.a). I allocated randomly 100 times the total number of chicks of each sex into the observed number of clutch sizes and obtained a final distribution which does not differ from that observed (G=8.291, df=9, p>0.05). I also tested whether the variance was greater than that expected under a random model, and found it to be within the expected one (p>0.05; Fig. 8.2.b).

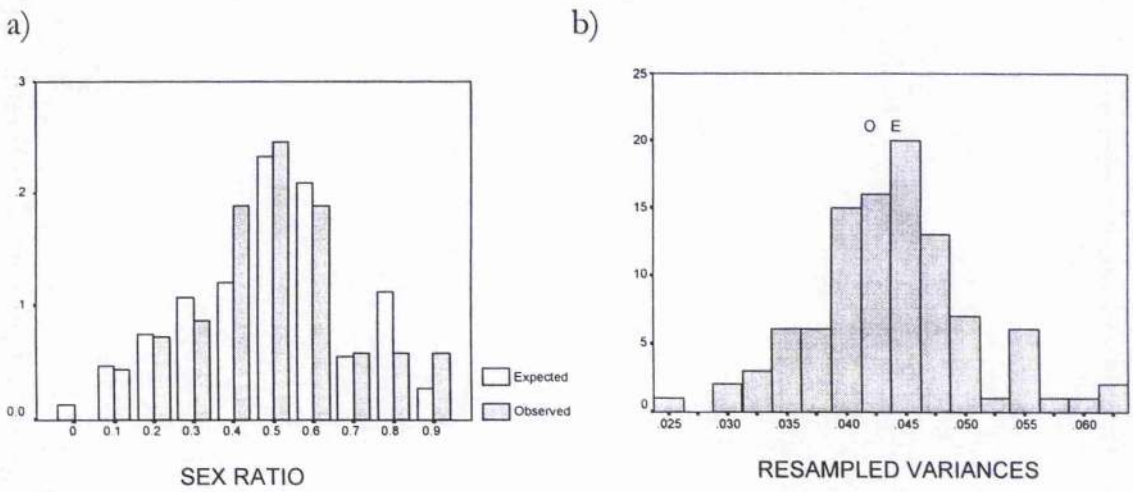


Figure 8.2. Observed (n=69 broods, brood sizes 4-8) and expected (resampled data: 100 simulations) sex ratio results. The two frequency distributions (a) do not differ significantly (G=8.291, df=9, p>0.05). The distribution of variances from the resampled data (b) shows that the observed variance in sex ratio (O) did not differ from that expected (E), (p>0.05).

If we look at the average sex ratio per clutch, there were no differences between years, ($F=1.16$, $df=3,55$, $P>0.05$) although there seems to be a tendency for sex ratios to alternate in direction between consecutive years (Fig. 8.3).

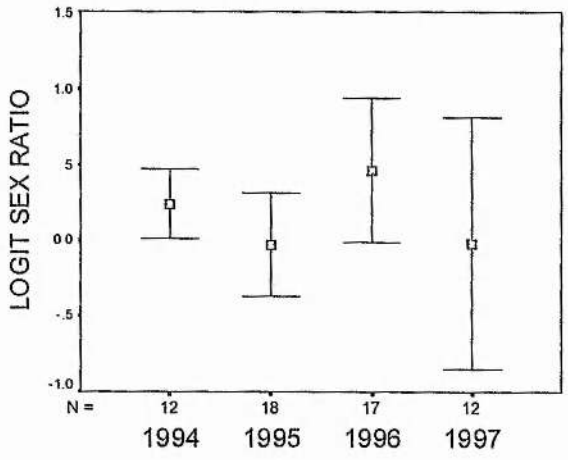


Figure 8.3. Average sex ratio for each of the years of the study ($F=1.16$, $df=3,55$, $P>0.05$).

The sample was divided into two sets, one consisting of independent data points, where only one clutch belonging to each male or female contributed to it, and a second one in which I knew the sex ratios of the clutch of the same female in two consecutive years. For the data set of independent clutches I conducted a general linear model analysis with several male characteristics as variables (Table 8.2).

explanatory variable	df	change in deviance	p
Repertoire size	1	3.73	0.053
Song length	1	2.07	0.150
Song versatility	1	0.26	0.612
Element rate	1	1.52	0.218
Repertoire fixation	1	0.03	0.859
Arrival date	1	0.05	0.822
Pairing date	1	0.19	0.662
Survival	1	0.89	0.344
Age	1	0.95	0.330

Table 8.2. Correlates of the proportion of males in the broods of willow warblers. The change in deviance has approximately a χ^2 distribution (Genstat 5 Committee 1993), and that is what the probability values correspond to. The sample size is 50 independent broods.

The results show that only repertoire size did account for a substantial change in deviance of the full model ($p=0.053$). The proportion of males in the brood decreases with increasing male song repertoire size (linear regression: $t=-2.41$, $n=50$, $p<0.016$; $\beta=-0.327$; Fig. 8.3).

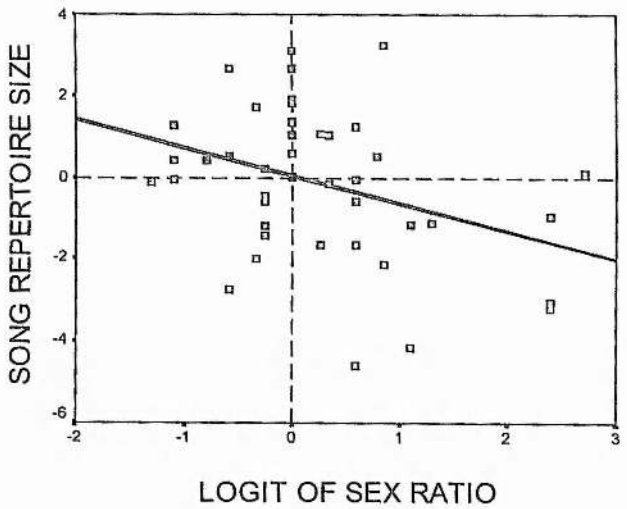


Figure 8.3. Proportion of males in a brood in relation to the repertoire size of the male. The regression line is: logit transformed sex ratio=-0.327+(0.135*repertoire size), (t=-2.41, n=50, p<0.016).

For a limited set of females data on their brood sex ratios could be obtained for two consecutive years. It is interesting to look at the repeatability of hatching sex ratio because it gives an indication of the degree of facultative manipulation that females can achieve. The repeatability value (Lessells & Boag 1987) for these 9 females is very low: r=-0.114. This could be due to an inherent female age effect or to an effect of the male on the sex ratio of the brood. I ran a two-way ANOVA with year and divorce as factors and the male characteristics as covariates. Neither of the factors is significant (year: F=1.637, df=1, p>0.05; divorce: F=0.05, df=1, p>0.05), suggesting that female age is not responsible for the change. However, the interaction between the two factors is significant (F=5.507, df=1, p<0.039). Females which pair again with the same male decrease the percentage of males in their broods, whereas those which change males, significantly increase it (Fig. 8.4).

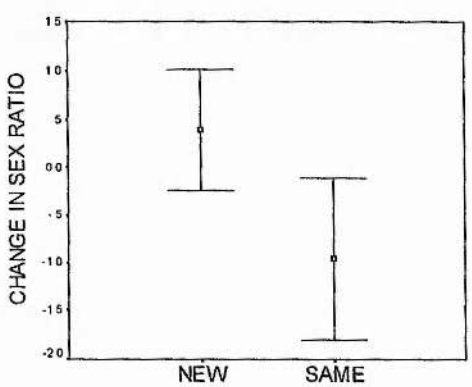


Figure 8.4. Change in sex ratio (logit transformed) for females which stay with their mates or change males in two consecutive years. Females seems to decrease the proportion of males in the brood it when remating with an old male (F=5.507, df=1,8, p<0.039).

From the 298 chicks sexed between 1994 and 1996, 48 (16%) were resighted or caught in the following year. There are sex differences in this respect, and males are more likely to recruit to the same population than females: 32 males (10.7%) and 16 females (5.36%) did so, this difference being significant on a G test ($G=5.436$, $df=1$, $p<0.05$). The sex ratio of a clutch does not correlate with the probability of recruiting young to the population ($r=-0.106$, $n=50$, $p>0.05$), nor did clutches with at least either one male, one female or one bird recruiting have different sex ratios ($t<0.6$, $n=50$, $p>0.05$, for the three tests).

It would be very interesting to know if the two sexes are equally costly to produce, for that would give us an idea about the possible benefits that biases in sex ratios may have. The only way I can approach this problem is by looking at differences in nestling size. However, a difference in size or weight does not automatically mean a difference in rearing cost. Evidence so far shows that differences in size or weight do not correlate with differences in cost or provisioning (Stamps 1990). For nine clutches, wing lengths and weights were measured in the chicks. As the age of the chicks was not the same in all clutches, I standardised values within clutches. Wing length did not differ between male and female nestlings ($t=-0,17$, $df=55$, $p>0.05$), but both weight and body condition (residual of the regression of weight on wing length) were significantly higher in males than in females (weight: $t=3.70$, $df=55$, $p<0.001$; body condition: $t=3.9$, $df=55$, $p<0.001$; Fig. 8.5).

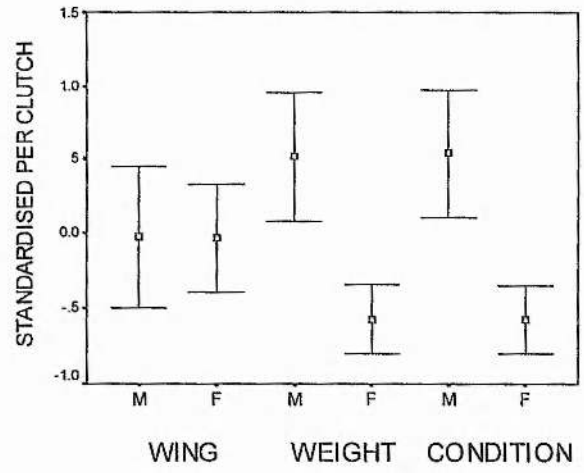


Figure 8.5. Sex differences in wing length, weight and body condition of nestling willow warblers. Wing length did not differ between male and female nestlings ($t=-0,17$, $df=55$, $p>0.05$), but both weight and body condition were significantly higher in males than in females (weight: $t=3.70$, $df=55$, $p<0.001$; body condition: $t=3.9$, $df=55$, $p<0.001$).

To check whether the degree of bias in the sex ratio of a given clutch influences the development of the rarer sex, I calculated the average wing, weight and body condition of

males and subtracted it from those of females. Unfortunately the variation in sex ratio present in this sample only ranges from female-biased to equality. However we can still test whether males do better in average when they are a minority in the nest or when the sex ratio is equal. The correlations show that males have shorter wings ($r=0.73$, $n=9$, $p<0.05$; Fig. 8.7a) and tend to have lighter weights ($r=0.62$, $n=9$, $p<0.082$; Fig. 8.7b) in female-biased clutches. Body condition also shows a similar trend but the correlation is not significant ($r=0.51$, $n=9$, $p>0.05$).

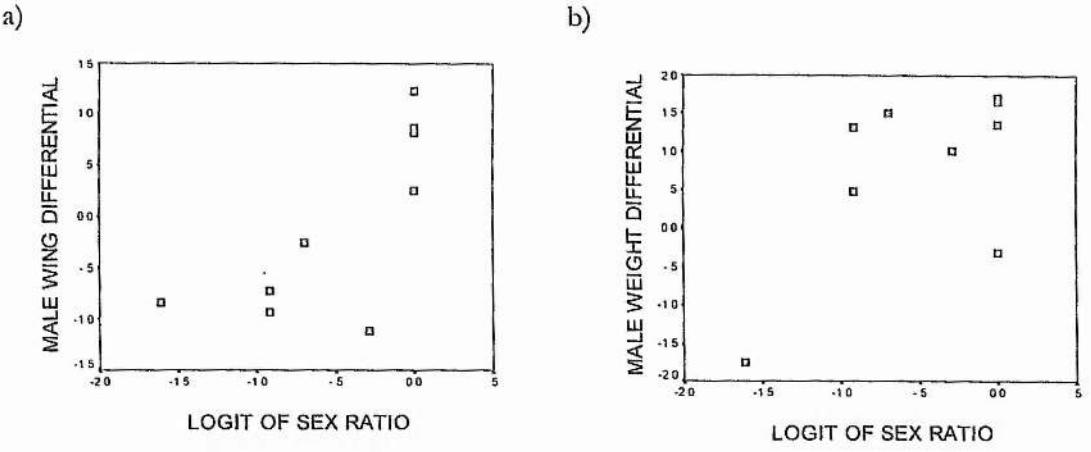


Figure 8.7. Males have lighter weights (b: $r=0.68$, $n=9$, $p<0.05$) and tend to have shorter wings (a: $r=0.61$, $n=9$, $p<0.082$) in female-biased than in even-sex clutches

Discussion

No discrepancies from Fisherian sex ratios were found in the studied population of willow warblers. This is in contrast to some recent studies which have found significant departures from binomial expectations in passerines, e.g. blue tit *Parus caeruleus* (Svensson & Nilsson 1996), great tit *Parus major* (Lessells et al. 1996), or great reed warbler *Acrocephalus arundinaceus* (Westerdahl et al. 1997). However, a 50:50 population wide sex ratio does not exclude the occurrence of adaptive individual manipulations (Williams 1979).

Of all male characteristics, only song repertoire size showed a sizeable contribution to the variation in sex ratio, and females paired to males with large repertoires had fewer sons than those paired to males with small repertoires. This is contrary to expectations (e.g. Burley 1981; Ellegren et al. 1996) based on the inheritance of characters which determine attractiveness or fitness. If reproductive success has higher variance in males, and females use male ornaments to assess quality in their partners, theory predicts that females should produce more sons than daughters when paired to highly ornamented males. In the case of the willow warbler, repertoire size does not explain patterns of pairing or extra-pair

paternity (Chapter 6 & 7). Repertoire size was however the only variable that predicted offspring survival, and survival in birds older than second years (Chapter 5).

One possible confounding variable is that of female age or breeding experience. Most adult females which return to breed pair with their previous male if he is available (Chapter 6). It is likely that there is a hidden correlation in the data set between female age and male song repertoire. The negative correlation between song repertoire and proportion of sons in the clutch could be due to older females producing an excess of females. Age effects on sex ratios have been found in the red-winged blackbird *Agelaius phoeniceus*, although in the opposite direction (Blank & Nolan 1983): older females produce more sons. By looking at the sample of females studied in two different years we can test if there is a consistent age effect. The data show that only females which mate again with the previous male increase the number of daughters. If there was a genuine female age/experience effect we would expect all females to show the same change. The fact that the change in sex ratio between years is correlated with the change in song versatility between the males, or the same male in two consecutive years, seems to give support to a sex ratio adjustment based on male characteristics. Why would a female produce a female-biased sex ratio when paired to a male with a large repertoire?

If females are less costly to produce, something which the sex difference in hatching weight may suggest, although this is not necessarily true (see Stamps 1990), it would seem that older females are reducing their breeding expenditure in broods raised later in life. This will not make evolutionary sense unless the probability of surviving increases with age for females, in which case it could pay females to reduce the breeding effort and maximise the number of years rather than the breeding output of each breeding attempt. This is opposite to what has been found in other passerines like the collared flycatcher, in which the probability of surviving does not vary with age for adults (Gustafsson & Pärt 1990), and where females increase their parental investment the older they are (Pärt et al. 1992).

One possible explanation for the negative correlation between male biased sex ratio and repertoire size is based on the sex based philopatry which this species shows, like most passerines (Clarke et al. 1997). Females disperse further, and the probability of a male chick returning to breed to the same population is double to that of a female (see Results). As viability is related to the expression of song repertoire size (Chapter 5), a female paired to a male of large repertoire should have higher chances of producing surviving offspring than one paired to a male with a small repertoire. If the probability of this offspring recruiting to the population is high enough, a female may benefit by adjusting the sex ratio of the brood

in order to reduce the amount of local competition between kin for mating opportunities (Hamilton 1967) or resources (Clark 1978). This could be achieved by moving the sex ratio away from the most competitive sex, in this case males (Bradbury et al. (1997) make a similar suggestion for the European starling *Sturnus vulgaris*). This suggestion fits with the evidence that females decrease the percentage of males in the clutch when pairing again with the same male.

One prediction from this hypothesis is that the probability of a male chick recruiting should be related to the sex ratio of the brood it has originated from. This prediction is not met by my results, but given the low survival rates among fledglings, the data set is probably too limited to test it. Body condition at fledging correlates with the probability of surviving in the collared flycatcher, another trans-Saharian migratory passerine (Lindén et al. 1992). If this also is the case in the willow warbler, the fact that males raised in female-biased broods tend to have worse body condition than their sisters suggests that female offspring are being favoured in female-biased broods. This may have repercussions on the viability of these chicks. Further work in the levels of begging and chick competition in clutches is necessary to elucidate how the rare sex fares in broods of biased sex ratio.

It is not possible to determine the adaptive nature of the pattern found without further experimental research. The 100% accuracy obtained in the sexing on the adult birds and the wide range of species to which this technique is applicable, makes the PCR procedure developed by Griffiths et al. (in press) the method of choice in studying sex ratio in bird broods. Results can be obtained in a very short time, and this would allow experimental manipulations in hatchling sex ratios. These experiments are necessary in order to test the validity of any adaptive hypothesis put forward to explain patterns of sex ratio variation in the wild.

Chapter 9

MALE RESPONSE TO DIFFERENT SONG REPERTOIRE SIZES

Abstract

A playback experiment was performed to test the hypothesis that large song repertoires elicit higher aggressive responses than small ones. Each experimental bird was tested with both small and large repertoires on different days. The results do not show an effect of either the type of stimulus or the order of presentation, but the interaction between these two factors is statistically significant. Males increased their response to the second playback if this one was the large repertoire, and decreased their response if it was the small one. It is suggested that the birds perceived the differences between stimuli, but the experimental design was not sensitive enough to detect a difference in the response.

Introduction

Sexual selection has been considered as a dual process ever since its first formulation (Darwin 1871). Darwin recognised that the evolution of exaggerated secondary sexual characters could be explained by two different mechanisms. The first consisted of competition "between individuals of the same sex, generally the males, in order to kill or drive away their rivals", and the second was that in which the exaggerated traits were used "to excite or charm those of the opposite sex, generally the females, which no longer remain passive, but select the most agreeable partners". These two modes are often referred to as "intrasexual" or "male-male competition", and "intersexual" or "female choice" respectively (Bradbury & Davies 1987; Andersson 1994).

The distinction between these two mechanisms is often blurred in nature. Females may choose males by assessing traits which are mainly used in male-male competition (West-Eberhard 1983), or even choose territory characteristics which are ultimately correlated with some male trait which they are not even able to perceive (Wiley & Poston 1996).

Further, the effects of sexual selection by male competition are not always easy to separate from those of natural selection. For instance, males with exaggerated secondary traits may enhance their probability of surviving by being better at obtaining resources outside the breeding season, without this directly affecting their mating success (Rohwer 1982).

Sexual selection by male contest competition faces fewer problems in its theoretical modelling than that based on mate choice (e.g. heritability of fitness is not a prerequisite), and the two mechanisms possess different evolutionary dynamics (Andersson 1994). Thus, assessing which one of the two mechanisms is operating in a certain case is a useful exercise in itself, which can help explain the evolution and maintenance of different mating systems (Bradbury & Davies 1987). A further problem is that signals used in both mechanism are likely to be the same. This is because only costly signals can evolved to be used in female choice or male-male contests, as only these good can signal male quality and are thus not amenable to cheating (Zahavi 1975; Grafen 1990).

The two main functions of birdsong, namely mate attraction and territorial defence (Darwin 1871; Catchpole & Slater 1995), fit nicely with the two modes of sexual selection. However, the evolutionary pressures brought about by these two mechanisms do not seem to follow universal directions, and can lead song characteristics to either simplicity or complexity depending on the species studied (Read & Weary 1992).

Tests of the effects of sexual selection by male-male competition should demonstrate that individuals with high values of the trait under study accrue enhanced mating success as a result of being better at competing for resources with other males. In the case of song repertoires in territorial passerines, a necessary proof of the role of sexual selection by male competition should provide evidence that males with larger repertoires are better at obtaining or defending territories than those with small ones. Ideally, the trait should be experimentally modified, but this it is not technically not possible as yet.

The best demonstration of these effects comes from experiments where territories are artificially occupied with loudspeakers broadcasting different repertoire sizes. Two experiments done on great tits *Parus major* (Krebs et al. 1978) and red-winged blackbirds *Agelaius phoeniceus* (Yasukawa 1981) have shown that those territories where repertoires are played back remain unoccupied for longer than those where single songs are used.

By contrast, it is not so straightforward to interpret experiments where the response of territorial males to different song repertoire sizes is measured (Searcy 1992). However, a case can be made for these experiments, because even if a more aggressive response is not functional per se, as Searcy argues, we can consider that differences in response reflect

different degrees of danger as perceived by the territorial male, and which are conveyed by the different repertoires. This could be so if repertoire size correlates with male quality or the male's resource-holding potential (Stoddard et al. 1988). There is some evidence now for this in several species (Lambrechts & Dhondt 1986; Hiebert et al. 1989; Horn et al. 1993; Lampe & Espmark 1994), as well as in the willow warbler (Chapter 3 & 5).

Several experiments have tested for differences in the response of territorial males to small and large repertoires (Lemon et al. 1981; Searcy 1983; Catchpole 1989) but most have failed to find a significant effect. However, it could be argued (Stoddard et al. 1988) that the playback designs used in most of these studies are not sensitive to the question being asked. This last article reports significant differences in the response of song sparrows *Melospiza melodia* to songtype variation when a more sensitive playback design is used. According to the authors, the main error of other studies was in measuring the immediate response of the bird, which is typically maximal, given the immediate impact of an intrusion. In this context, all response variables show a ceiling effect, which obscures any motivational difference that the stimuli may bring about in a less aggressive situation. Stoddard et al. (1988) avoided this by conducting the playback from outside the territories, and by looking at the long-term response rather the immediate one.

The present experiment uses this approach to test the response of territorial willow warblers to small and large repertoires. The working assumption is that if birds with large repertoires are better quality males, territorial birds should show higher and more sustained levels of aggression towards the playback of large repertoires than towards that of a single songtype.

Methods

Subjects were 16 male willow warblers holding territories in two small mixed woods of birch, willow and spruce in East Fife (Kippo Wood and Redwells Wood), Scotland. The approximate total populations of these woods are 80 and 60 territorial males respectively. Experiments were performed between 0900 and 1200 hours between 27 May and 4 June 1997. The exact breeding status of 10 of the males was known, and all of these were paired and their female was incubating. This was probably the case for all subjects, as during these dates all known females in the more extensive breeding study in Kippo Wood had already begun to incubate. The precise limits of the territories were known prior to the experiment. These territories had been stable for at least two weeks, the time at which the last arrival had squeezed into the area.

The rationale of the experimental design comes mainly from Stoddard et al. (1988). The intention was to simulate as mild as possible an intrusion, so that ceiling effects in the response were avoided. The experiment did not start until the females were no longer fertile, and was thus after the period during which males show their highest aggression towards playback (Arvidsson 1992). The loudspeaker was placed 15 m outside the territory, on the ground. To avoid conflicts with other neighbours, only territories which bordered barley or rape plantations were used.

Before the full playback was presented, birds were attracted to the territory boundary by playing songs from the loudspeaker at a rate of one song per minute. The rationale of this warming-up period was to make sure that all birds were within hearing range and at the same distance from the loudspeaker. All subjects came to the desired location in less than 4 min. During this period none of them received less than 2 songs or more than 4 (average 2.8, SD=0.34), so it seems reasonable to think that this caused no bias in the results. Once the bird was within sight, the playback started. The only response variable measured was the number of songs per minute that the bird sang in response. This variable was recorded during three periods of observation: during the 3 min that the playback lasted (DURING), a first post-experimental period of 5 min (POST-1), and further period of 5 min (POST-2).

A total of 8 sets of playback stimuli were used. These were recorded in the previous two years from 8 males with large repertoires that were no longer resident in the area and which none of the experimental subjects had had as neighbours. Each set of stimuli was prepared by editing good song recordings from the subject, and classifying the songs into clusters looking at the elements that the songs had in common (see Chapter 1 for a detailed description of the technique). I will use the term songtype instead of cluster for convenience. For the different stimuli, very defined clusters of song variants were chosen, so that different songtypes only share one or two elements with each other. Each set of stimuli was composed of a SINGLE SONGTYPE stimulus, which was composed of random presentations of 8 different recordings of one songtype; and a REPERTOIRE stimulus, which was also composed of random songs from 2 different recordings from each of 4 songtypes (see Fig. 9.1 for an example). The REPERTOIRE stimulus always included as one of the songs that used as the SINGLE SONGTYPE stimulus. Note that both stimuli were composed of the same number of different recordings, so that the effect of any possible background effects was minimised. Songs were edited with SIGNAL and modified so that they all had the same average volume. Each stimulus track was 3-min long and was recorded on different tracks on a DAT machine.

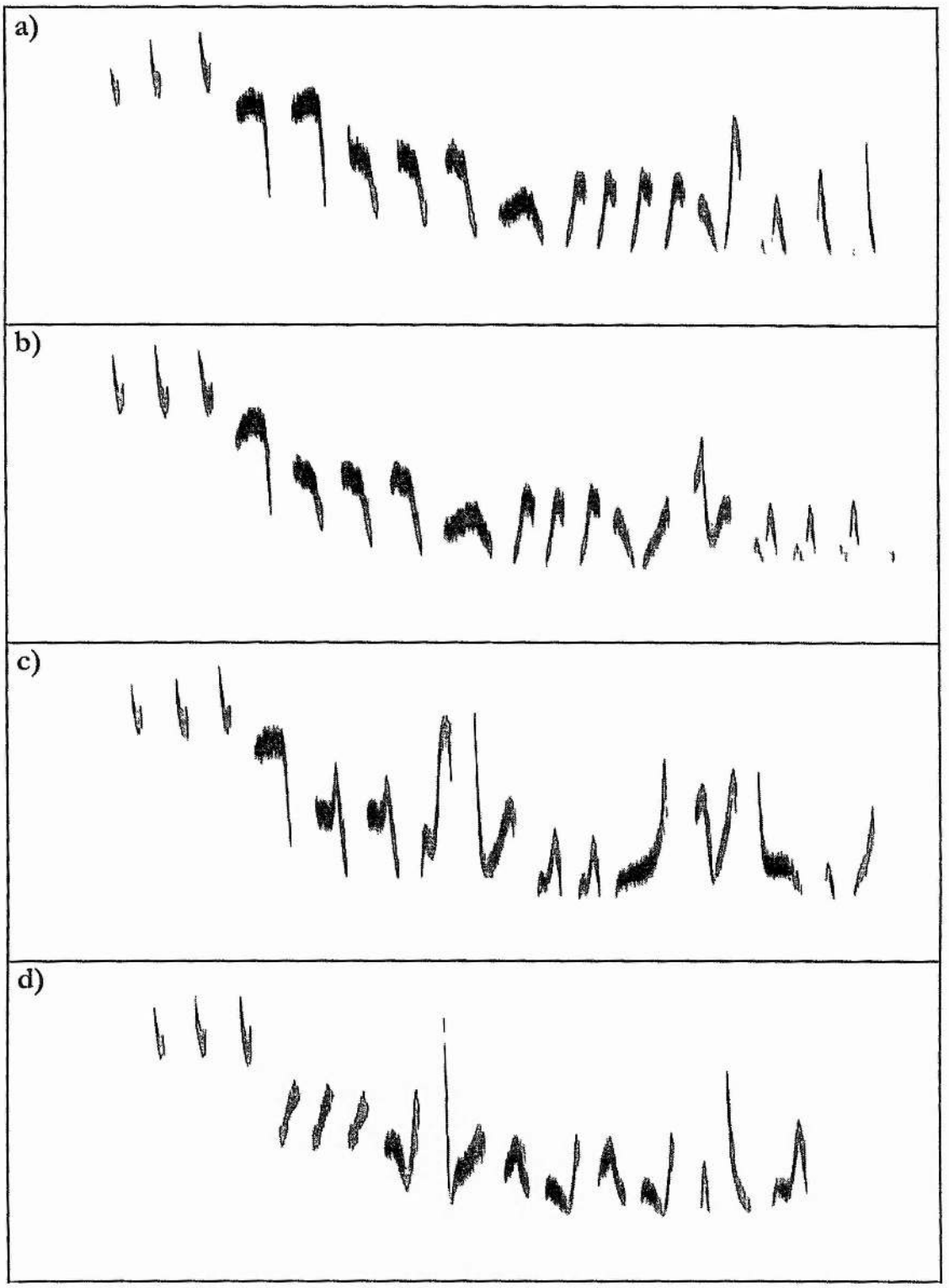


Figure 9.1. Sonograms of four songs used as stimuli in the experiment. They are all from the same individual. Songs a and b correspond to the same songtype, whereas c and d are representatives of other two different songtypes. The x axis represents 3 seconds, and the y axis runs from 1 kHz to 8 kHz.

The experiment was carried out using a SONY DAT recorder TCD-D8, and a NAGRA DSM MONITOR loudspeaker. The volume was the same in all trials, and was adjusted to sound a bit louder than what a willow warbler sounds like in the field to an observer. A coaxial cable of 20 m connected the loudspeaker to the tape recorder, allowing the experimenter to control the bird's response from that distance.

Each of the 16 birds received a single set of stimuli, in a balanced order so that half of them were tested with the REPERTOIRE stimulus first, and the other half with the SINGLE SONGTYPE first. Each set of stimuli was used twice, but the order of presentation was different in the two occasions. Following Kroodsma's recommendations (1990) for the design of playback experiments, a single-way ANOVA was run for the response variable in each of the two conditions and periods, to test for possible spurious effects of any of the tapes. None of these analyses were significant (all $Fs>1.21$, $df=7,1$), so it was concluded that each of the two experiments done with each set of tapes could be pooled together in a Wilcoxon signed-rank test. This approach was preferred to Kroodsma's suggestion of treating each set of stimuli as a single data point. Pseudoreplication was not considered an issue in the post-hoc ANOVA done to test for effects of order of presentation and interaction, because only one bird or stimulus was present in the groups created after splitting the data into factors.

Results

All birds responded to the warming-up playback by approaching the border of the territory and displaying. Typically the bird twitches the tail and wings, which are held slightly unfolded, while it appears to look around for the source of the sounds, turning the body sideways (May 1949; Marchetti 1993). Flights between branches and trees are common, but often difficult to quantify, and thus were not used as a measure of response. In two cases the response also involved intensive high-pitched calling, and extreme levels of arousal. The playback in these two cases was stopped, as the aim of the experiment was to provoke moderate responses. One of these birds was seen close to its female, which suggests that the high response was due to the female being in the fertile period (Arvidsson 1992).

While the playback lasted, most birds did not sing much, though some did. This response seems to reflect an individual style, which may be related to the particular stage of breeding or other individual variables. Birds showed high consistency in how much they sang during playback regardless of the stimulus being presented. The correlation between song

responses for the same bird in the two trials was very high (r_s = 0.810, n=16, p < 0.01; Fig. 9.2). The singing response normally began after playback stopped. The quality of the songs also changed, though no recording were made that could allow objective comparisons. Songs produced during or immediately after playback were fainter and shorter, with high-pitched calls interspersed amongst them, as previously described (Järvi et al. 1980).

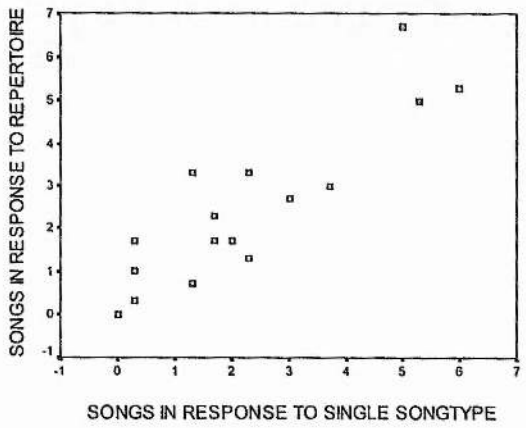


Figure 9.2. Intra-individual correlation between the number of songs produced in the two experimental trials.

The experimental design was aimed at comparing the response of individuals between the two different experimental conditions. There were no significant differences in any of the three stages of the playback (Wilcoxon signed-rank tests: z=-0.806, ns, during playback; z=- 1.450, ns, in the first post-experimental period; and z=-1.580, ns, in the second post-experimental period; n=16 for all tests). The average response to the single songtype was always lower than to the song repertoire, and this difference tended to increase with time from the playback, although it was never statistically significant (Fig. 9.3).

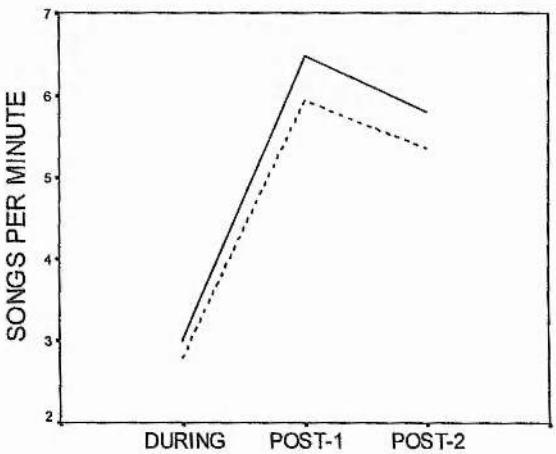


Figure 9.3. Average number of songs in response to the two experimental conditions. Upper line is the response to the repertoire, and the lower one the response to the single songtype.

An analysis of variance was run to test for possible effects of order of presentation. Here, the power of intra-subject comparisons disappears, but it is still a useful analysis to test for habituation and sensitisation effects. The results confirm those found before, that there were no differences of response to the two experimental conditions (see Table 9.1). No significant effects of order of presentation were found either, but the interaction of the two factors were significant in both post-experimental periods: birds exposed first to the repertoire stimulus, decreased their response when subsequently tested with the single songtype; those tested first with the single songtype increase their response when the repertoire was presented in the second session (Fig 9.4).

	Repertoire size	Order of presentation	Interaction
DURING	0.110 ns	0.198 ns	2.154 ns
POST-1	1.271 ns	0.248 ns	10.061 p<0.01
POST-2	0.724 ns	0.002 ns	3.981 p<0.05

Table 9.1. Results of the two-way ANOVA run separately for the variable number of songs produced in response at the three different stages of the playback. F and significance values are shown, df = (1, 28) in all cases.

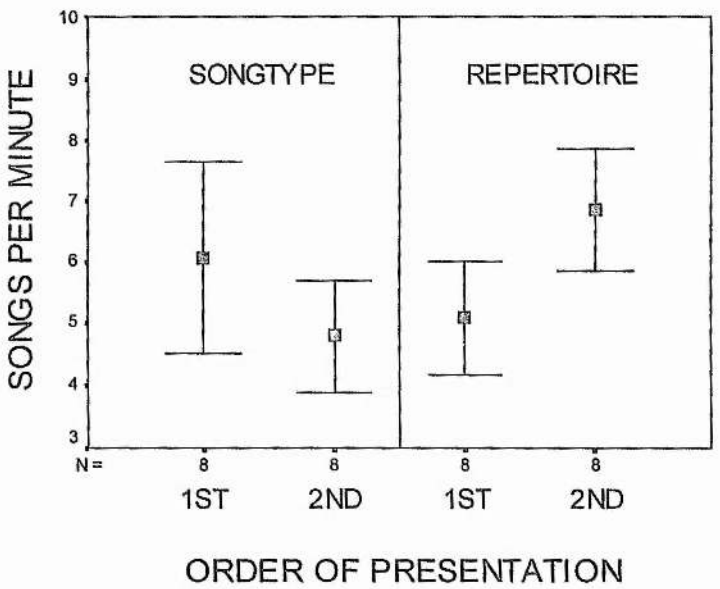


Figure 9.4. Average number of songs (±2SE) in response to the two experimental conditions, depending on the position in the order of presentation. See how there are no differences in response between stimuli or order of presentation, but an interaction between the two.

Discussion

The results did not support the hypothesis that repertoires elicit greater aggressive response than single songtypes in territorial willow warblers. However, there is an interesting pattern consistent with the predictions if the data are analysed taking into account the long-term response of the subjects. The interaction of the response with the order of presentation of the playback stimuli suggests that the subjects did react differently to the repertoire, although this could not be picked up by the experimental design.

Playback experiments conducted on territorial males do not always allow a straightforward interpretation (Searcy 1992). The fact that repertoires seem to reduce habituation, as the results suggest, cannot be seen as a function of song in that context, as there would seem to be no advantage to intruders to induce higher levels of aggression. As Searcy et al. (Searcy et al. 1995) have pointed out: "eliciting aggressive responses cannot be considered a function of song". The response must be placed in the context where it occurs. A playback simulates an attempt at settling a territory in the neighbourhood of other male's territory. (Jakobsson 1988) has shown that birds that have held a territory for more than one season regain their territory from a new settler if they are removed for some days from their territories, whereas first year males do not manage to do so. Given that this experiment has been carried out when all birds were already paired, we would expect the subjects to be extremely defensive over their territory and to show high persistence in its defence, given the pay-off asymmetry between owner and intruder (Maynard Smith & Parker 1976; Leimar & Enquist 1984). In four years of observations, I have never observed a take-over by a new male after females have arrived and paired up. Under these circumstances, it seems reasonable to expect the territorial male to defend his territory by all means regardless of who is the intruder. The failure of many studies to show discrimination between repertoire and single songtype in this situation could well be the consequence of this ceiling effect.

It makes sense to think of a short playback experiment as a take-over attempt followed by a complete retreat. It should be stressed that the lack of song after the playback is also a signal. The territorial bird is faced with the kind of information that it would get if an intruder had given up a take over attempt after having been challenged by him. In a sense it is a victory for the territory owner, and his motivational state must reflect this delusion. The fact that the bird encounters the same intrusion again a day later and responds differently depending on the repertoire size seems to imply that this variable conveys some information about the quality of the male or its motivation to fight. We know that

individual willow warblers sing with similar amounts of immediate variety (see Chapter 1), so this information is probably static (Gerhardt 1992) and typical of the individual, rather than context dependent. The increased response found may thus be functional if large repertoires signal higher Resource Holding Potential, and therefore pose a higher threat. This is very likely in this species, as repertoire size increases with age (Chapter 3) and is a good predictor of survival in birds older than two-years (Chapter 5). A relationship between male quality (e.g. dominance, territory size, age) and repertoire has been found in several other species (Lambrechts & Dhondt 1986; Hiebert et al. 1989; Horn et al. 1993; Lampe & Espmark 1994).

A remarkable result was the intra-individual correlation between the responses to the two playbacks. It suggests that variation between individuals in response might account for the greater part of the variance in response. This means that much larger sample of experimental males would be necessary in order to tests for differences in response. This variation in response is likely to be individual, given that all birds were at the same stage of the breeding stage. Strength of response has been used as an indication of male quality, and it could be argued that those birds that responded with more songs were of higher quality. However, song is only one component of response, and such an experiment should quantify other behaviours like flights or attack to a dummy.

Five years have past; five summers, with the length
Of five long winters (...)

W. WORDSWORTH

CONCLUSIONS

THE TRAIT

Song and singing

The multivariate analysis of the song of the willow warbler revealed five principal components: 1) repertoire size, 2) song length, 3) song versatility, 4) element rate; and 5) repertoire fixation. Repertoire size is the most variable song characteristic between individuals, and is highly repeatable within individuals in the same year. There are no songtypes as such, but sequences of elements with rather high probabilities of transition among them. The organisation of these sequences forms a hierarchically branching pattern. Although the element repertoire can be large, some elements appear very often, and there is a limited core of song sequences which birds use much more frequently than others. The singing style suggests that consecutive songs are drawn at random from the repertoire of element sequences.

Repeatability and heritability

Heritabilities could not be calculated due to the small sample size available and the confounding factor of extra-pair fertilisations. However, repeatabilities for several song characteristics (repertoire size, song length and versatility) were both significant and substantial (>0.4). Given a minimum heritability, the song characteristic which would have the highest value of additive genetic variance is repertoire size.

CORRELATES

Age

In the analysis of age and song, both cross-sectional and longitudinal approaches showed that several song characteristics change with age. These were: repertoire size, element rate and song versatility. Changes occur mainly between the second and third year, with no important development from that age on. In terms of the development of sexual characteristics, this can be seen as evidence that the song is not fully developed until the bird is in its third year.

Arrival date

Arrival date has been shown to be an important component of fitness in several species of migratory passerines. In this population of willow warblers, older males arrived earlier, and two song characteristics were found to predict arrival date: repertoire size and song length. Birds that arrived earlier had larger repertoires and shorter songs.

Male survival

The song characteristics that predict male survival were different for second years and older birds. Individuals with long songs in their second year were less likely to survive than those with short ones. In older birds, however, survival was positively correlated with song repertoire and element rate. The general tendency is for all these three variables to change the sign of the regression with survival in the two age classes.

Reproductive success

Repertoire size was the only song characteristic that predicted the number of young fledging from a male's nest. Although repertoire size did not correlate with number of recruits, males with large repertoires were more likely to have at least one recruit in the population than males with short repertoires.

FEMALE CHOICE

No song characteristic explained female pairing patterns consistently. In some years some variables had some predictive value, but the effect was not consistent between years. Considering only those males that arrived before the first female, there was a very high correlation between pairing date and male arrival date. This can be taken to imply that females are choosing territory quality rather than males. Extra-pair paternity was common in the population, with 43% of nests having at least one extra-pair nestling. Cuckolded males had significantly shorter songs, but there were no differences in repertoire size or any other measured variable. To the extent that females control extra-pair fertilisations, this constitutes a measure of mate choice.

MALE-MALE COMPETITION

I tested the effect of song repertoire size in the response of territorial males to playback. It was predicted that if repertoire size signalled male quality, males should respond more strongly or for longer to the playback of a repertoire than to a single songtype. There were

no differences between the response to these two conditions. However, there was an interaction between the treatment and the order of presentation of playback. This suggests that, although the experimental design failed to pick up differences in response, birds were somehow able to perceive differences between the two treatments.

BEHAVIOUR DURING THE FERTILE PERIOD

Mate guarding in this species was characterised by a pronounced drop in singing. Song did not seem to act as a paternity guard, and intrusions by other males were not related to the song rate of the territory owner. Song rate was not higher when intrusions occurred. Probably, the most effective paternity guard is close contact. The evidence runs counter the predictions of the fertility announcement hypothesis (Møller 1991).

WHAT DID NOT WORK...

In three consecutive years I tried to record the contribution of males to the feeding of the chicks. I tried hides (handmade bush style and commercial ones), video recording (handycams and black-and-white cell cams) and an automatic detection system (my own brainchild) with different magnets on the rings of male and female. None of these worked. The reasons for this are manifold, but they come down to the wariness of males when feeding the chicks. One expects evolution to select very strongly any behaviour that avoids nest predation, so it is not a surprise that feedings at the nest are so difficult to watch.

GENERAL DISCUSSION

The succinct rationale of the research done in this thesis is the following paradox: assuming that repertoires are costly to produce and store (e.g. Canady et al. 1984), why do males have repertoires rather than a single song? A possible reason is that, if only males of good quality or in good condition are able to produce these repertoires, good quality males would outweigh the costs of the repertoire by increased benefits in mating success (Zahavi 1975; Sutherland & De Jong 1991). This is because females pay a higher reproductive cost than males in most organisms. Thus, to the extent that male quality influences offspring viability (via direct benefits or good genes), females should be more discriminating than males when choosing mates (Andersson 1994). In the case of monogamous birds, variance in male

mating success can come about by differential female fertility, differential female parental effort, or through extra-pair paternity (Price et al. 1988; Birkhead & Møller 1992).

Male age was found to correlate with repertoire size in this study. If we only consider males which have completely developed their song (i.e. older than second-years), there is a positive correlation between repertoire size and survival. The probability of a male having an offspring recruiting in the population was also correlated with repertoire size. Thus, there is some evidence that repertoire size correlates with male quality. Repertoire size was repeatable (0.49) once the song was fully developed. Although this does not mean that the trait is heritable, the large phenotypic variance of the trait suggests that, given a minimum heritability, substantial levels of additive genetic variance may exist for this trait (Houle 1992).

Number of fledglings was positively correlated with repertoire size, even when the effect of arrival date was taken into account. As there was no nestling mortality in the population during the period of the study, number of fledglings is probably a very close correlate of clutch size, i.e. females laid more eggs when mated to males with large repertoires. However, female choice bore no relation to repertoire size. Females paired first with early arriving males. It seems reasonable to conclude that other variables of male quality which I did not measure, like song rate, may have influenced this choice. Females did not use repertoire size either when choosing extra-pair partners. Another characteristic, song length, seemed to be behind this choice, and males with short songs were more likely to be cuckolded. The evidence for indirect selection of repertoire size is thus scarce and conflicting.

The effects of song repertoires in male-male competition were examined by means of a playback experiment. The results did not provide conclusive evidence of an effect of repertoire in male-male competition. However, this kind of experiment is not the ideal way of testing for effects of repertoire size in male-male competition. Speaker replacement experiments in empty territories have been the best approach so far (e.g. Krebs et al. 1978).

Taken together, the evidence gathered in this thesis suggests that, although repertoire size correlates with several measures of male quality, the maintenance of this trait does not seem to be based in a disproportionally higher mating benefit. Without knowing the cost of repertoire size, it is not possible to assess to what extent repertoires are maladaptive. Besides, repertoires may confer additional benefits not measured here. For instance, large repertoires may allow males to match neighbours' songs (Slater 1981). It has been found in

other species that birds that share songs have higher reproductive success than those that do not (Beletsky & Orians 1989). Beecher et al. (1994) suggested that repertoire sharing, rather than repertoire size, was the target of sexual selection in the song sparrow. Although they did not specify whether this selection would come about by male-male competition or by mate choice, as song sharing reflects breeding experience and previous territory occupancy, it could play an important role in both mechanisms. Males prospecting for territories could use this trait to assess the likelihood that the owner would defend the territory (experiments have shown that previous occupancy is an important factor in contests over territories: Jakobsson 1988), and females could make use of this information to choose experienced, successful males.

An interesting possibility for the evolution of repertoire size is suggested by the low repeatability of repertoire size when we consider second-years in the calculation. Although the repeatability of the fully developed trait is higher, what this effectively means is that a female could not use this trait as a reliable indicator of the bird's 'genetic' or final repertoire, unless she can assess the male's age. Similar scenarios have been found for other traits, like badges in sparrows (Veiga 1995) and spurs in pheasants (Grahn & von Schantz 1994). This might be the reflection of a hidden trade-off between current reproduction and survival (sensu Williams 1966). If the costs of the expression of a sexual character are negatively correlated with the time taken to develop them, individuals of low quality could accelerate the development of a trait at the expense of survival possibilities. If this trait is sexually selected, the benefits of the current year's reproductive success may render this strategy adaptive. If this is the case, a survival pattern like that found in Fig. 5.5 is to be expected.

Further research in this species should answer two important questions that were not tackled in this thesis: the relationship between song rate and repertoire size, and the quality of male paternal care in relation to song characteristics. Another interesting issue is whether females are able to control EPCs, and what the costs of these might be. Ideally, experimental manipulation of EPP should be attempted. A fascinating topic which remains to be analysed in the tapes I have recorded in Kippo is the phenomenon of song sharing. Males that have been neighbours for long share a large proportion of their repertoires, much more than young males new to the area. This might indicate the presence of a constraint that would render song sharing an honest signal of breeding experience and previous territory occupancy. The interface of this function with song learning strategies and constraints seems worth exploring.

REFERENCES

Aebischer, A., Perrin, N., Krieg, M., Studer, J. & Meyer, D. R. 1996. The role of territory choice, mate choice and arrival date on the breeding success in the Savi's warbler *Locustella luscinoides*. *Journal of Avian Biology* **27**, 143-152.

Alatalo, R. V., Höglund, J. & Lundberg, A. 1991. Lekking in the black grouse: a test of male viability. *Nature* **352**, 155-156.

Alatalo, R. V. & Lundberg, A. 1986. The sexy son hypothesis: data from the pied flycatcher *Ficedula hypoleuca*. *Animal Behaviour* **34**, 1454-1462.

Alatalo, R. V., Lundberg, A. & Glynn, C. 1986. Female pied flycatchers choose territory quality and not male characteristics. *Nature* **323**, 152-153.

Alatalo, R. V., Lundberg, A. & Stahlbrandt, K. 1984. Female mate choice in the pied flycatcher *Ficedula hypoleuca*. *Behavioral Ecology and Sociobiology* **14**, 253-261.

Alatalo, R. V., Mappes, J. & Elgar, M. A. 1997. Heritabilities and paradigm shifts. *Nature* **385**, 402-403.

Andersson, M. 1994. *Sexual Selection*. Princeton: Princeton University Press.

Appleby, B. M., Petty, S. J., Blakey, J. K., Rainey, P. & Macdonald, D. W. 1997. Does variation of sex ratio enhance reproductive success of offspring in tawny owls (*Strix aluco*). *Proceedings of the Royal Society of London, Series B* **264**, 1111-1116.

Arnold, S. J. 1994. Multivariate inheritance and evolution: a review of concepts. In: *Quantitative Genetic Studies of Behavioural Evolution*. (Ed. by C. R. B. Boake), pp. 17-48. Chicago: Chicago University Press.

Arvidsson, B. & Klaesson, P. 1986. Territory size in a willow warbler *Phylloscopus trochilus* population in mountain birch forest in Swedish Lapland. *Ornis Scandinavica* **17**, 24-30.

Arvidsson, B. L. 1992. Copulation and mate guarding in the willow warbler. *Animal Behaviour* **43**, 501-510.

Arvidsson, B. L. & Neergard, R. 1991. Mate choice in the willow warbler: a field experiment. *Behavioral Ecology and Sociobiology* **29**, 225--229.

Bakker, T. C. M. 1993. Positive genetic correlation between female preference and preferred male ornament in sticklebacks. *Nature* **363**, 255-257.

Barber, C. A., Robertson, R. J. & Boag, P. T. 1996. The high frequency of extra-pair paternity in tree swallows is not an artefact of nestboxes. *Behavioral Ecology and Sociobiology* **38**, 425-430.

Beecher, M. D., Campbell, S. E. & Stoddard, P. K. 1994. Correlation of song learning and territory establishment strategies in the song sparrow. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 1450-1454.

Beletsky, L. D. & Orians, G. H. 1989. Familiar neighbors enhance breeding success in birds. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 7933-7936.

Bensch, S. & Hasselquist, D. 1992. Evidence for active female choice in a polygynous warbler. *Animal Behaviour* **44**, 301-311.

Bensch, S., Hasselquist, D. & von Schantz, T. 1994. Genetic similarity between parents predicts hatching failure: non-incestous inbreeding in the great reed warbler? *Evolution* **48**, 317-326.

Bensch, S., Price, T. & Kohn, J. 1997. Isolation and characterization of microsatellite loci in a *Phylloscopus* warbler. *Molecular Ecology* **6**, 91-92.

Berthold, P. 1993. *Bird Migration: A General Survey.* Oxford: Oxford University Press.

Birkhead, T. R. & Biggins, J. D. 1987. Reproductive synchrony and extra-pair copulations in birds. *Ethology* **74**, 320-334.

Birkhead, T. R., Buchanan, K. L., Devoogd, T. J., Pellatt, E. J., Székely, T. & Catchpole, C. K. 1997. Song, sperm quality and testes asymmetry in the sedge warbler. *Animal Behaviour* **53**, 965-971.

Birkhead, T. R. & Møller, A. P. 1992. *Sperm Competition in Birds: Evolutionary Causes and Consequences.* London: Academic Press.

Bjørnstad, G. & Lifjeld, J. T. 1996. Male parental care promotes early fledging in an open-nester, the willow warbler *Phylloscopus trochilus*. *Ibis* **138**, 229-235.

Bjørnstad, G. & Lifjeld, J. T. 1997. High frequency of extra-pair paternity in a dense and synchronous population of willow warblers *Phylloscopus trochilus*. *Journal of Avian Biology* **28**, 319-324.

Blank, J. L. & Nolan, V. 1983. Offspring sex ratio in red-winged blackbirds is dependent on maternal age. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 6141-6145.

Boake, C. R. B. 1989. Repeatability: its role in evolutionary studies of mating behavior. *Evolutionary Ecology* **3**, 173-182.

Bradbury, J. W. & Davies, N. B. 1987. Relative roles of intra- and intersexual selection. In: *Sexual Selection: Testing the Alternatives.* (Ed. by J. W. Bradbury & M. B. Andersson), pp. 143-163. Chichester: Wiley.

Bradbury, R. B., Cotton, P. A., Wright, J. & Griffiths, R. 1997. Nestling sex ratio in the European starling *Sturnus vulgaris*. *Journal of Avian Biology* **28**, 255-258.

Buchanan, K. L. & Catchpole, C. K. 1997. Female choice in the sedge warbler, *Acrocephalus schoenobaenus*: multiple cues from song and territory quality. *Proceedings of the Royal Society of London, Series B* **264**, 521-526.

Burley, N. 1981. Sex ratio manipulation and selection for attractiveness. *Science* **211**, 721-722.

Burley, N. 1982. Facultative sex ratio manipulation. *American Naturalist* **120**, 81-107.

Burley, N. 1986. Sex-ratio manipulation in color-banded populations of zebra finches. *Evolution* **40**, 1191-1206.

Burley, N. 1988. The differential-allocation hypothesis: an experimental test. *American Naturalist* **132**, 611-628.

Burley, N. T., Parker, P. G. & Lundy, K. 1996. Sexual selection and extrapair fertilization in a socially monogamous passerine, the zebra finch (Taeniopygia guttata). *Behavioral Ecology* **7**, 218-226.

Butlin, R. K. & Hewitt, G. M. 1986. Heritability estimates for characters under sexual selection in the grasshopper, *Chorthippus brunneus*. *Animal Behaviour* **34**, 1256-1261.

Canady, R. A., Kroodsma, D. E. & Nottebohm, F. 1984. Population differences in complexity of a learned skill are correlated with the brain space involved. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 6232-6234.

Catchpole, C., Leisler, B. & Dittami, J. 1986. Sexual differences in the responses of captive great reed warblers (*Acrocephalus arundinaceus*) to variation in song structure and repertoire size. *Ethology* **73**, 69-77.

Catchpole, C. K. 1976. Temporal and sequential organisation of song in the sedge warbler (*Acrocephalus arundinaceus*). *Behaviour* **59**, 226-245.

Catchpole, C. K. 1980. Sexual selection and the evolution of complex songs among European warblers of the genus *Acrocephalus*. *Behaviour* **74**, 149-165.

Catchpole, C. K. 1982. The evolution of bird sounds in relation to mating and spacing behaviour. In: *Acoustic Communication in Birds. Vol. 1. Production, Perception and Design Features of Sounds.* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 253-295. New York: Academic Press.

Catchpole, C. K. 1986. Song repertoires and reproductive success in the great reed warbler *Acrocephalus arundinaceus*. *Behavioral Ecology and Sociobiology* **19**, 439-445.

Catchpole, C. K. 1989. Responses of male sedge warblers to playback of different repertoire sizes. *Animal Behaviour* **37**, 1046-1047.

Catchpole, C. K. & Slater, P. J. B. 1995. *Bird Song: Biological Themes and Variations.* Cambridge: Cambridge University Press.

Charnov, E. L. 1982. *The Theory of Sex Allocation.* Princeton: Princeton University Press.

Cheverud, J. M. & Moore, A. J. 1994. Quantitative genetic and the role of the environment provided by relatives in behavioural evolution. In: *Quantitative Genetic Studies of Behavioral Evolution.* (Ed. by C. R. B. Boake), pp. 67-100. Chicago: University of Chicago Press.

Clark, A. B. 1978. Sex ratio and local resource competition in a prosimian primate. *Science* **201**, 163-165.

Clarke, A. L., Saether, B.-E. & Røskaft, E. 1997. Sex biases in avian dispersal: a reappraisal. *Oikos* **79**, 429-438.

Clutton-Brock, T. H. 1991. *The Evolution of Parental Care.* Princeton: Princeton University Press.

Colegrave, N., Birkhead, T. R. & Lessells, C. M. 1995. Sperm precedence in zebra finches does not require special mechanisms of sperm competition. *Proceedings of the Royal Society of London, Series B* **259**, 223-228.

Coltman, D. W., Bowen, W. D. & Wright, J. M. 1988. Birth weight and neonatal survival of harbour seal pups are positively correlated with genetic variation measured by microsatellites. *Philosophical Transactions of the Royal Society of London, Series B* **265**, 803-809.

Cosens, S. E. & Sealy, S. G. 1986. Age-related variation in song repertoire sharing of yellow warblers (*Dendroica petechia*). *Canadian Journal of Zoology* **64**, 1926-1929.

Coulson, T. N., Pemberton, J. M., Albon, S. D., Beaumont, M., Marshall, T. C., Slate, J., Guiness, F. E. & Clutton-Brock, T. H. 1998. Microsatellites reveal heterosis in red deer. *Proceedings of the Royal Society of London, Series B* **265**, 489-495.

Currie, D. R., Burke, T., Whitney, R. L. & Thompson, D. B. A. 1998. Male and female behaviour and extra-pair paternity in the wheatear. *Animal Behaviour* **55**, 689-703.

Cuthill, I. C. & Macdonald, W. A. 1990. Experimental manipulation of the dawn and dusk chorus in the blackbird *Turdus merula*. *Behavioral Ecology and Sociobiology* **26**, 209-216.

Daan, S., Dijkstra, C. & Weissing, F. J. 1996. An evolutionary explanation for seasonal trends in avian sex ratios. *Behavioral Ecology* **7**, 426-430.

Dale, S., Rinden, H. & Slagsvold, T. 1992. Competition for a mate restricts mate search of female pied flycatchers. *Behavioral Ecology and Sociobiology* **30**, 165-176.

Dale, S. & Slagsvold, T. 1996. Mate choice on multiple cues, decision rules and sampling strategies in female pied flycatchers. *Behaviour* **133**, 903-944.

Darwin, C. 1871. *The Descent of Man, and Selection in Relation to Sex.* London: Murray.

Davies, N. B. 1992. *Dunnock Behaviour and Social Evolution.* Oxford: Oxford University Press.

Dawkins, R. & Dawkins, M. 1973. Decisions and the uncertainty of behaviour. *Behaviour* **45**, 83-103.

Dawkins, R. & Dawkins, M. 1974. Some descriptive and explanatory stochastic models of decision-making. In: *Motivational control system analysis.* (Ed. by D. J. McFarland), pp. 119-168. London: Academic Press.

De Lope, F. & Møller, A. P. 1993. Female reproductive effort depends on the degree of ornamentation of their mates. *Evolution* **47**, 1152-1160.

Devoogd, T. J., Krebs, J. R., Healy, S. D. & Purvis, A. 1993. Relations between song repertoire size and the volume of brain nuclei related to song: comparative evolutionary analyses amongst oscine birds. *Proceedings of the Royal Society of London, Series B* **254**, 75-82.

Dickinson, J. L. 1997. Male detention affects extra-pair copulation frequency and pair behaviour in western bluebirds. *Animal Behaviour* **53**, 561-571.

Dunn, P. O. & Robertson, R. J. 1993. Extra-pair paternity in polygynous tree swallows. *Animal Behaviour* **45**, 231-239.

Dunn, P. O., Whittingham, L. A., Lifjeld, J. T., Robertson, R. J. & Boag, P. T. 1994. Effects of breeding density, synchrony, and experience on extrapair paternity in tree swallows. *Behavioral Ecology* **5**, 123-129.

Eason, P. 1992. Optimization of territory shape in heterogeneous habitat: a field study of the red-capped cardinal (*Paroaria gularis*). *Journal of Animal Ecology* **61**, 411-424.

Eason, P. & Hannon, S. J. 1994. New birds on the block: new neighbors increase defensive costs for territorial male willow ptarmigan. *Behavioral Ecology and Sociobiology* **34**, 419-426.

Eason, P. K. & Stamps, J. A. 1992. The effect of visibility on territory size and shape. *Behavioral Ecology* **3**, 166-172.

Eeens, M., Pinxten, R. & Verheyen, R. F. 1992. No overlap in song repertoire size between yearling and older starlings *Sturnus vulgaris*. *Ibis* **134**, 72-76.

Ellegren, H. 1992. Polymerase-Chain-Reaction (PCR) analysis of microsatellites: a new approach to studies of genetic relationships in birds. *Auk* **109**, 886-895.

Ellegren, H., Gustafsson, L. & Sheldon, B. C. 1996. Sex ratio adjustment in relation to paternal attractiveness in a wild bird population. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 11723-11728.

Eriksson, D. & Wallin, L. 1986. Male bird song attracts females: a field experiment. *Behavioral Ecology and Sociobiology* **19**, 297-299.

Espmark, Y. O. & Lampe, H. M. 1993. Variations in the song of the pied flycatcher within and between breeding seasons. *Bioacoustics* **5**, 33-65.

Falconer, D. S. & Mackay, T. F. C. 1996. *Introduction to Quantitative Genetics*. Harlow: Longman.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. *Department of Genetics, University of Washington, Seattle* ,

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

Fitzpatrick, S. 1994. Colourful migratory birds: evidence for a mechanism other than parasite resistance for the maintenance of 'good genes' sexual selection. *Proceedings of the Royal Society of London, Series B* **257**, 155-160.

Freeman-Gallant, C. R. 1997. Extra-pair paternity in monogamous and polygynous Savannah sparrows, *Passerculus sanwichensis*. *Animal Behaviour* **53**, 397-404.

Frey, D. F. & Pimentel, R. A. 1978. Principal component analysis and factor analysis. In: *Quantitative Ethology*. (Ed. by P. W. Colgan), pp. 219-246. New York: John Wiley & Sons.

Fridolfsson, A. K., Gyllensten, U. B. & Jakobsson, S. 1997. Microsatellite markers for paternity testing in the willow warbler *Phylloscopus trochilus*: high frequency of extra-pair young in an island population. *Hereditas* **126**, 127--132.

Genstat 5 Committee. 1993. *Genstat 5 Release 3 Reference Manual*. Oxford: Oxford University Press.

Gerhardt, H. C. 1992. Multiple messages in acoustic signals. *Seminars in the Neurosciences* **4**, 391-400.

Gibbs, H. L., Weatherhead, P. J., Boag, P. T., White, B. N., Tabak, L. M. & Hoysak, D. J. 1990. Realized reproductive success of polygynous red-winged blackbirds revealed by DNA markers. *Science* **250**, 1394-1397.

Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L. & Feldman, M. W. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463-471.

Gomendio, M., Clutton-Brock, T. H., Albon, S. D., Guiness, F. E. & Simpson, M. J. 1990. Mammalian sex ratios and variation in costs of rearing sons and daughters. *Nature* **343**, 261-263.

Gowaty, P. A. & Bridges, W. C. 1991. Nestbox availability affects extra-pair fertilizations and conspecific nest parasitism in eastern bluebirds, *Sialia sialis*. *Animal Behaviour* **41**, 661-675.

Grafen, A. 1990. Sexual selection unhandicapped by the Fisher process. *Journal of theoretical Biology* **144**, 473-516.

Grahn, M. 1993. Mortality in the pheasant *Phasianus colchicus* during the breeding season. *Behavioral Ecology and Sociobiology* **32**, 95-101.

Grahn, M. & von Schantz, T. 1994. Fashion and age in pheasants: age differences in mate choice. *Proceedings of the Royal Society of London, Series B* **255**, 237-241.

Graves, J., Ortega-Ruano, J. & Slater, P. J. B. 1993. Extra-pair copulations and paternity in shags: do females choose better males? *Proceedings of the Royal Society of London, Series B* **253**, 3-7.

Gray, E. M. 1997. Female red-winged blackbirds accrue material benefits from copulating with extra-pair males. *Animal Behaviour* **53**, 625-639.

Greig-Smith, P. W. 1982. Seasonal patterns of song production by male Stonechats *Saxicola torquata*. *Ornis Scandinavica* **13**, 225-231.

Griffiths, R., Daan, S. & Dijkstra, C. 1996. Sex identification in birds using two CHD genes. *Proceedings of the Royal Society of London, Series B* **263**, 1251-1256.

Griffiths, R., Double, M., Orr, K. & Dawson, R. in press. A simple DNA test to sex most birds. *Molecular Ecology* ,

Guilford, T. & Dawkins, M. S. 1991. Receiver psychology and the evolution of animal signals. *Animal Behaviour* **42**, 1-14.

Gustafsson, L. 1986. Lifetime reproductive success and heritability: empirical suport for Fisher's fundamental theorem. *American Naturalist* **128**, 761-764.

Gustafsson, L. & Pärt, T. 1990. Acceleration of senescence in the collared flycatcher *Ficedula albicollis* by reproductive costs. *Nature* **347**, 279-281.

Gyllensten, U. B., Jakobsson, S. & Temrin, H. 1990. No evidence for illegitimate young in monogamous and polygynous warblers. *Nature* **343**, 168-170.

Halliday, T. & Arnold, S. J. 1987. Multiple mating by females: a perspective from quantitative genetics. *Animal Behaviour* **35**, 939-941.

Hamilton, W. D. 1967. Extraordinary sex ratios. *Science* **156**, 477-488.

Hartshorne, C. 1973. *Born to Sing*. Bloomington: Indiana University Press.

Hasselquist, D. & Bensch, S. 1991. Trade-off between mate guarding and mate attraction in the polygynous great reed warbler. *Behavioral Ecology and Sociobiology* **28**, 187-193.

Hasselquist, D., Bensch, S. & von Schantz, T. 1996. Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature* **381**, 229-232.

Helb, H.-W. 1973. Analyse der artisolierenden Parameter im Gesang des Fitis (*Phylloscopus t. trochilus*) mit Untersuchungen zur Objektivierung der abalytischen Methode. *Journal für Ornithologie* **114**, 145-206.

Hewison, A. J. M. & Gaillard, J. M. 1996. Birth sex ratios and local resource competition in roe deer, *Capreolus capreolus*. *Behavioral Ecology* **7**, 461-464.

Hiebert, S. M., Stoddard, P. K. & Arcese, P. 1989. Repertoire size, territory acquisition and reproductive success in the song sparrow. *Animal Behaviour* **37**, 266-273.

Hooff, J. A. R. A. M. v. 1982. Categories and sequences of behavior: methods of description and analysis. In: *Handbook of methods in nonverbal behavior research*. (Ed. by K. R. Scherer & P. Ekman), pp. Cambridge: Cambridge University Press.

Horn, A. G., Dickinson, T. E. & Falls, J. B. 1993. Male quality and song repertoires in western meadowlarks (*Sturnella neglecta*). *Canadian Journal of Zoology* **71**, 1059-1061.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* **130**, 195-204.

Houtman, A. M. 1992. Female zebra finches choose extra-pair copulations with genetically attractive males. *Proceedings of the Royal Society of London, Series B* **249**, 3-6.

Howard, R. D. 1974. The influence of sexual selection and interspecific competition on mockingbird song (*Mimus polyglottos*). *Evolution* **28**, 428-438.

Jacobs, L. F. 1996. Sexual selection and the brain. *Trends in Ecology and Evolution* **11**, 82-86.

Jakobsson, S. 1988. Territorial fidelity of willow warbler (*Phylloscopus trochilus*) males and success in competition over territories. *Behavioral Ecology and Sociobiology* **22**, 79-84.

Järvi, T. 1983. The evolution of song versatility in the willow warbler *Phylloscopus trochilus*: a case of evolution by intersexual selection explained by the "female's choice of the best mate". *Ornis Scandinavica* **14**, 123-128.

Järvi, T., Radesäter, T. & Jakobsson, S. 1980. The song of the willow warbler *Phylloscopus trochilus* with special reference to singing behaviour in agonistic situations. *Ornis Scandinavica* **11**, 236-242.

Jeffreys, A. J., Allen, M. J., Hagelberg, E. & Sonnberg, A. 1992. Identification of the skeletal remains of Josef Mengele by DNA analysis. *Forensic Science International* **56**, 65-76.

Jenni, L. & Winkler, R. 1994. *Moult and Ageing of European Passerines*. London: Academic Press.

Ji, L., Macaubas, C., Hallmayer, J., Kimura, A. & Mignot, E. 1996. Mutation rate varies among alleles at a microsatellite locus: phylogenetic evidence. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 15285-15288.

Johnson, L. S. & Kermott, L. H. 1991. The functions of song in male house wrens (*Troglodytes aedon*). *Behaviour* **116**, 190-209.

Kempenaers, B. 1997. Does reproductive synchrony limit male opportunities or enhance female choice for extra-pair paternity? *Behaviour* **134**, 551-562.

Kempenaers, B., Verheyen, G. R. & Dhont, A. A. 1995. Mate guarding and copulation behaviour in monogamous and polygynous blue tits: do males follow a best-of-a-bad-job strategy? *Behavioral Ecology and Sociobiology* **36**, 33-42.

Kempenaers, B., Verheyen, G. R. & Dhont, A. A. 1997. Extrapair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics, and offspring quality. *Behavioral Ecology* **8**, 481-492.

Kempenaers, B., Verheyen, G. R., van den Broeck, M., Burke, T., van Broeckhoven, C. & Dhont, A. A. 1992. Extra-pair paternity results from female preference for high-quality males in the blue tit. *Nature* **357**, 494-496.

Kirkpatrick, M. & Ryan, M. J. 1991. The evolution of mating preferences and the paradox of the lek. *Nature* **350**, 33-38.

Komdeur, J., Daan, S., Tinbergen, J. & Mateman, C. 1997. Extreme adaptive modification in sex ratio of the Seychelles warbler's eggs. *Nature* **385**, 522-525.

Krackow, S. 1995. The developmental asynchrony hypothesis for sex ratio manipulation. *Journal of theoretical Biology* **176**, 273-280.

Krebs, J. R., Ashcroft, R. & Webber, M. 1978. Song repertoires and territory defence in the Great Tit. *Nature* **271**, 539-542.

Krokene, C., Anthonisen, K., Lifjeld, J. T. & Amundsen, T. 1996. Paternity and paternity assurance behaviour in the bluethroat *Luscinia s. svecica*. *Animal Behaviour* **52**, 405-417.

Kroodsma, D. E. 1990. Using appropriate experimental designs for intended hypotheses in 'song' playbacks, with examples for testing effects of song repertoire sizes. *Animal Behaviour* **40**, 1138-1150.

Kroodsma, D. E. & Canady, R. A. 1985. Differences in repertoire size, singing behaviour, and associated neuroanatomy among marsh wren populations have a genetic basis. *Auk* **102**, 439-446.

Kruskal, J. B. 1983. An overview of sequence comparison. In: *Time Warps, String Edits, and Macromolecules: the Theory and Practice of Sequence Comparison.* (Ed. by D. Sankoff & J. B. Kruskal), pp. 1-44. Reading: Addison-Wesley.

Lambrechts, M. & Dhondt, A. A. 1986. Male quality, reproduction and survival in the great tit (*Parus major*). *Behavioral Ecology and Sociobiology* **19**, 57-63.

Lampe, H. M. & Espmark, Y. O. 1994. Song structure reflects male quality in pied flycatchers, *Ficedula hypoleuca*. *Animal Behaviour* **47**, 869-876.

Leimar, O. & Enquist, M. 1984. Effects of asymmetries in owner-intruder conflicts. *Journal of theoretical Biology* **111**, 475-491.

Lemon, R. E. & Chatfield, C. 1971. Organization of song in cardinals. *Animal Behaviour* **19**, 1-17.

Lemon, R. E., Fieldes, M. A. & Struger, J. 1981. Testing the monotony threshold hypothesis of bird song. *Zeitschrift für Tierpsychologie* **56**, 359-379.

Lessells, C. M. & Boag, P. T. 1987. Unrepeatable repeatabilities: a common mistake. *Auk* **104**, 116-121.

Lessells, C. M., Mateman, A. C. & Visser, J. 1996. Great tit hatchling sex ratios. *Journal of Avian Biology* **27**, 135-143.

Lijfeld, J. T., Dunn, P. O., Robertson, R. J. & Boag, P. T. 1993. Extra-pair paternity in monogamous tree swallows. *Animal Behaviour* **45**, 213-229.

Lindén, M., Gustafsson, L. & Pärt, T. 1992. Selection on fledging mass in the collared flycatcher and the great tit. *Evolution* **73**, 336-343.

Lundberg, A. & Alatalo, R. V. 1992. *The Pied Flycatcher*. London: T & AD Poyser.

Manning, J. T. 1985. Choosy females and correlates of male age. *Journal of theoretical Biology* **116**, 349-354.

Manning, J. T. 1989. Age-advertisement and the evolution of the peacock's train. *Journal of Evolutionary Biology* **2**, 379-384.

Marchetti, K. 1993. Dark habitats and bright birds illustrate the role of the environment in species divergence. *Nature* **362**, 149-152.

Marchetti, K. & Price, T. 1989. Differences in the foraging of juvenile and adult birds: the importance of developmental constraints. *Biological Reviews* **64**, 51-70.

Marler, P. & Peters, S. 1981. Sparrows learn adult song and more from memory. *Science* **213**, 780-782.

Marler, P. & Peters, S. 1982. Long-term storage of learned birdsongs prior to production. *Animal Behaviour* **30**, 479-482.

Martin, P. & Bateson, P. 1993. *Measuring Behaviour: An Introductory Guide.* Cambridge: Cambridge University Press.

May, D. J. 1949. Studies on a community of willow warblers. *Ibis* **91**, 24-54.

Maynard Smith, J. 1987. Sexual selection: a classification of models. In: *Sexual Selection: Testing the Alternatives.* (Ed. by J. W. Bradbury & M. B. Andersson), pp. 9-20. New York: John Wiley.

Maynard Smith, J. & Parker, G. A. 1976. The logic of asymmetric contests. *Animal Behaviour* **24**, 159-175.

McDonald, M. V. 1989. Function of song in Scott's seaside sparrow *Ammodramus maritimus peninsulae. Animal Behaviour* **38**, 468-485.

McGregor, P. K., Krebs, J. R. & Perrins, C. M. 1981. Song repertoires and lifetime reproductive success in the great tit (*Parus major*). *American Naturalist* **118**, 149-159.

McKinney, F., Derrickson, S. R. & Mineau, P. 1983. Forced copulation in waterfowl. *Behaviour* **86**, 250-294.

Merilä, J. & Sorjonen, J. 1994. Seasonal and diurnal patterns of singing and song-flight activity in bluethroats (*Luscinia svecica*). *Auk* **111**, 556-562.

Møller, A. P. 1988. Spatial and temporal distribution of song in the yellowhammer *Emberiza citrinella. Ethology* **78**, 321-331.

Møller, A. P. 1991. Preferred males acquire mates of higher phenotypic quality. *Proceedings of the Royal Society of London, Series B* **245**, 178-182.

Møller, A. P. 1991. Viability in positively related to degree of ornamentation in male swallows. *Proceedings of the Royal Society of London, Series B* **243**, 145-148.

Møller, A. P. 1991. Why mated songbirds sing so much: mate guarding and male announcement of mate fertility status. *American Naturalist* **138**, 994-1014.

Møller, A. P. 1994. Male ornament size as a reliable cue to enhanced offspring viability in the barn swallow. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 6929-6932.

Møller, A. P. 1994. Phenotype-dependent arrival time and its consequences in a migratory bird. *Behavioral Ecology and Sociobiology* **35**, 115-122.

Møller, A. P. 1996. The cost of secondary sexual characters and the evolution of cost-reducing traits. *Ibis* **138**, 112-119.

Møller, A. P. & Pomiankowski, A. 1993. Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology* **32**, 167-176.

Møller, A. P. & Tegelström, H. 1997. Extra-pair paternity and tail ornamentation in the barn swallow *Hirundo rustica. Behavioral Ecology and Sociobiology* **41**, 353-360.

Mountjoy, D. J. & Lemon, R. E. 1996. Female choice for complex song in the European starling: a field experiment. *Behavioral Ecology and Sociobiology* **38**, 65-71.

Muller, K. L., Stamps, J. A., Krishnan, V. V. & Willits, N. H. 1997. The effects of conspecific attraction and habitat quality on habitat selection in territorial birds (*Troglodytes aedon*). *American Naturalist* **150**, 650-661.

Mundinger, P. C. 1995. Behaviour-genetic analysis of canary song: inter-strain differences in sensory learning, and epigenetic rules. *Animal Behaviour* **50**, 1491-1511.

Nelson, D. A. 1992. Song overproduction and selective attrition lead to song sharing in the field sparrow (Spizella pusilla). *Behavioral Ecology and Sociobiology* **30**, 415-424.

Nemeth, K. 1996. Different singing styles in mated and unmated reed buntings *Emberiza schoeniclus. Ibis* **138**, 172-176.

Neudorf, D. L., Stutchbury, B. J. M. & Piper, W. H. 1997. Covert extraterritorial behavior of female hooded warblers. *Behavioral Ecology* **8**, 595-600.

Norris, K. 1993. Heritable variation in a plumage indicator of viability in male great tits *Parus major*. *Nature* **362**, 537-539.

Nottebohm, F. & Nottebohm, M. E. 1978. Relationship between song repertoire and age in the canary *Serinus canaria*. *Zeitschrift für Tierpsychologie* **46**, 298-305.

Nowicki, S. & Podos, J. 1993. Complexity, coupling, and contingency in the production of birdsong. In: *Perspectives in Ethology, Vol. 10: Behavior and Evolution*. (Ed. by P. P. G. Bateson, P. H. Klopfer & N. S. Thompson), pp. New York: Plenum Press.

Oddie, K. 1998. Sex discrimination before birth. *Trends in Ecology and Evolution* **13**, 130-131.

Omland, K. E. 1996. Female mallard preferences for multiple male ornaments: I: Natural variation. *Behavioral Ecology and Sociobiology* **39**, 353-360.

Orell, M., Koivula, K., Rytkonen, S. & Lahti, K. 1994. To breed or not to breed: causes and implications of non-breeding habit in the willow tit *Parus montanus*. *Oecologia* **100**, 339-346.

Otter, K., Chruszcz, B. & Ratcliffe, L. 1997. Honest advertisement and song output during the dawn chorus of black-capped chickadees. *Behavioral Ecology* **8**, 167-173.

Pärt, T. 1991. Is dawn singing related to paternity insurance? The case of the collared flycatcher. *Animal Behaviour* **41**, 451-456.

Pärt, T., Gustafsson, L. & Moreno, J. 1992. "Terminal investment" and a sexual conflict in the collared flycatcher (*Ficedula albicollis*). *American Naturalist* **140**, 868-882.

Perreault, S., Lemon, R. E. & Kuhnlein, U. 1997. Patterns and correlates of extrapair paternity in American redstarts (*Setophaga ruticilla*). *Behavioral Ecology* **8**, 612-621.

Petrie, M. 1994. Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature* **371**, 598-599.

Petrie, M. & Kempenaers, B. 1998. Extra-pair paternity in birds: explaining variation between species and populations. *Trends in Ecology and Evolution* **13**, 52-58.

Podos, J., Peters, S., Rudnicky, T., Marler, P. & Nowicki, S. 1992. The organization of song repertoires in song sparrows: themes and variations. *Ethology* **90**, 89-106.

Pomiankowski, A. & Møller, A. P. 1995. A resolution of the lek paradox. *Proceedings of the Royal Society of London, Series B* **260**, 21-29.

Price, T., Kirkpatrick, M. & Arnold, S. J. 1988. Directional selection and the evolution of breeding date in birds. *Science* **240**, 798-799.

Radesäter, T., Jakobsson, S., Andbjer, N., Bylin, A. & Nyström, K. 1987. Song rate and pair formation in the willow warbler *Phylloscopus trochilus*. *Animal Behaviour* **35**, 1645-1651.

Rätti, O., Hovi, M., Lundberg, A., Tegelström, H. & Alatalo, R. V. 1995. Extra-pair paternity and male characteristics in the pied flycatcher. *Behavioral Ecology and Sociobiology* **37**, 419-425.

Read, A. F. & Weary, D. M. 1992. The evolution of bird song: comparative analyses. *Philosophical Transactions of the Royal Society of London, Series B* **338**, 165-187.

Reid, M. L. & Weatherhead, P. J. 1990. Mate choice criteria of Ipswich sparrows: the importance of variability. *Animal Behaviour* **40**, 538-544.

Rodrigues, M. 1996. Song activity in the chiffchaff: territorial defence or mate guarding? *Animal Behaviour* **51**, 709-716.

Rohwer, S. 1982. The evolution of reliable and unreliable badges of fighting ability. *American Zoologist* **22**, 531-546.

Saino, N., Primmer, C. R., Ellegren, H. & Møller, A. P. 1997. An experimental study of paternity and tail ornamentation in the barn swallow (*Hirundo rustica*). *Evolution* **51**, 562-570.

Schluter, D. 1988. Estimating the form of natural selection on a quantitative trait. *Evolution* **47**, 658-667.

Schubert, M. 1967. Probleme der Motivwahl und der Gesangsaktivität bei *Phylloscopus trochilus* (L.). *Journal für Ornithologie* **108**, 265-194.

Schubert, M. 1976. Über die Variabilität von Lockrufen des Gimpels *Pyrrhula pyrrhula*. *Ardea* **64**, 62-71.

Schwabl, H. 1993. Yolk is a source of maternal testosterone for developing birds. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 11446-11450.

Schwabl, H. 1996. Environment modifies the testosterone levels of a female bird and its eggs. *Journal of Experimental Zoology* **276**, 157-163.

Schwabl, H. 1996. Maternal testosterone in the avian egg enhances postnatal growth. *Comparative Biochemistry and Physiology* **114A**, 271-276.

Schwabl, H., Mock, D. W. & Gieg, J. A. 1997. A hormonal mechanism for parental favouritism. *Nature* **386**, 231.

Searcy, W. A. 1983. Response to multiple song types in male song sparrows and field sparrows. *Animal Behaviour* **31**, 948-949.

Searcy, W. A. 1984. Song repertoire size and female preferences in song sparrows. *Behavioral Ecology and Sociobiology* **14**, 281-286.

Searcy, W. A. 1992. Measuring responses of female birds to male song. In: *Playback and studies of animal communication.* (Ed. by P. K. McGregor), pp. 175-189. New York: Plenum Press.

Searcy, W. A. 1992. Song repertoire and mate choice in birds. *American Zoologist* **32**, 71-80.

Searcy, W. A. & Andersson, M. 1986. Sexual selection and the evolution of song. *Annual Review of Ecology and Systematics* **17**, 507-533.

Searcy, W. A. & Marler, P. 1981. A test for responsiveness to song structure and programming in song sparrows. *Science* **213**, 926-928.

Searcy, W. A., Podos, J., Peters, S. & Nowicki, S. 1995. Discrimination of song types and variants in song sparrows. *Animal Behaviour* **49**, 1219-1226.

Sheldon, B. C. 1994. Song rate and paternity in the chaffinch. *Animal Behaviour* **47**, 986-987.

Sheldon, B. C. & Ellegren, H. 1996. Offspring sex and paternity in the collared flycatcher. *Proceedings of the Royal Society of London, Series B* **263**, 1017-1021.

Sheldon, B. C., Merilä, J., Qvarnström, A., Gustafsson, L. & Ellegren, H. 1997. Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. *Proceedings of the Royal Society of London, Series B* **264**, 297-302.

Siegel, S. & Castellan, N. J. 1988. *Nonparametric Statistics.* New York: McGraw Hill.

Slagsvold, T. 1996. Dawn and dusk singing of male American robins in relation to female behaviour. *Wilson Bulletin* **108**, 507-515.

Slater, P. J. B. 1973. Describing sequences of behaviour. In: *Perspectives in ethology.* (Ed. by P. P. G. Bateson & P. H. Klopfer), pp. 131-153. New York: Plenum.

Slater, P. J. B. 1978. Data collection. In: *Quantitative Ethology*. (Ed. by P. W. Colgan), pp. 7-24. New York: John Wiley & Sons.

Slater, P. J. B. 1981. Chaffinch song repertoires: observations, experiments and a discussion of their significance. *Zeitschrift für Tierpsychologie* **56**, 1-24.

Slater, P. J. B. 1983. Sequences of song in chaffinches. *Animal Behaviour* **31**, 272-281.

Sokal, R. R. & Rohlf, F. J. 1995. *Biometry*. New York: Freeman.

Soukup, S. S. & Thompson, C. F. 1997. Social mating system affects the frequency of extra-pair paternity in house wrens. *Animal Behaviour* **54**, 1089-1105.

Stamps, J. A. 1990. When should avian parents differentially provision sons and daughters? *American Naturalist* **135**, 671-685.

Stamps, J. A. 1991. The effect of conspecifics on habitat selection in territorial species. *Behavioral Ecology and Sociobiology* **28**, 29-36.

Steinberg, J. B. 1977. Information theory as an ethological tool. In: *Quantitative methods in the study of animal behavior*. (Ed. by B. A. Hazlett), pp. 47-74. New York: Academic Press.

Stoddard, P. K., Beecher, M. D. & Willis, M. S. 1988. Response of territorial male song sparrows to song types and variations. *Behavioral Ecology and Sociobiology* **22**, 125-130.

Stresemann, E. & Stresemann, V. 1966. Die Mauser der Vögel. *Journal für Ornithologie, Sonderheft.* ,

Stutchbury, B. J. & Morton, E. S. 1995. The effect of breeding synchrony on extra-pair mating systems in songbirds. *Behaviour* **132**, 675-690.

Sullivan, M. S. 1994. Mate choice as an information gathering process under time constraint: implications for behaviour and signal design. *Animal Behaviour* **47**, 141-151.

Sundberg, J. 1992. Absence of mate guarding in the yellowhammer (*Emberiza citrinella*). *Ethology* **92**, 242-256.

Sundberg, J. & Dixon, A. 1996. Old, colourful male yellowhammers, *Emberiza citrinella*, benefit from extra-pair copulations. *Animal Behaviour* **52**, 113-122.

Sutherland, W. J. & De Jong, M. C. M. 1991. The evolutionary stable strategy for secondary sexual characters. *Behavioral Ecology* **2**, 16-19.

Svensson, E. 1997. Natural selection on avian breeding time: causality, fecundity-dependent, and fecundity-independent selection. *Evolution* **51**, 1276-1283.

Svensson, E. & Nilsson, J.-Å. 1996. Mate quality affects offspring sex ratio in blue tits. *Proceedings of the Royal Society of London, Series B* **263**, 357-361.

Svensson, L. 1984. *Identification Guide to European Passerines*. Stockholm: Published by the author.

Thielcke, G. & Linsenmair, K. E. 1963. Zur geographischen Variation des Gesanges des Zilpzalps, *Phylloscopus collybita*, in Mittel- und Südwesteuropa mit einem Vergleich des Gesanges des Fitis, *Phylloscopus trochilus*. *Journal für Ornithologie* **104**, 372-402.

Titus, R. C., Chnadler, C. R., Ketterson, E. D. & Nolan, V. 1997. Song rates of dark-eyed juncos do not increase when females are fertile. *Behavioral Ecology and Sociobiology* **41**, 165-169.

Todt, D. 1970. Zur Ordnung im Gesang der Nachtigall (*Luscinia megarhynchos* L.). *Verhandlungen der Deutschen Zoologischen Gesellschaft* **64**, 249-252.

Todt, D. & Hultsch, H. 1996. Acquisition and performance of song repertoires: ways of coping with diversity and versatility. In: *Ecology and Evolution of Acoustic Communication in Birds*. (Ed. by D. E. Kroodsma & E. H. Miller), pp. 79-96. Ithaca: Cornell University Press.

Trivers, R. L. 1972. Parental investment and sexual selection. In: *Sexual Selection and the Descent of Man.* (Ed. by B. Campbell), pp. 136-179. Chicago: Aldine Publishing Company.

Trivers, R. L. & Willard, D. E. 1973. Natural selection of parental ability to vary the sex ratio of offspring. *Science* **179**, 90-92.

Underhill, L. G., Prys-Jones, R. P., Dowsett, R. J., Herroelen, P., Johnson, D. N., Lawn, M. R., Norman, S. C., Pearson, D. J. & Tree, A. J. 1992. The biannual primary moult of Willow Warblers *Phylloscopus trochilus* in Europe and Africa. *Ibis* **134**, 286-297.

Valdes, A. M., Slatkin, M. & Freimer, N. B. 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* **133**, 737-749.

Vallet, E., Kreutzer, M. & Beme, I. 1998. Two-notes syllables in the canary songs elicit high levels of sexual display. *Animal Behaviour* **55**, 291-297.

Veiga, J. P. 1993. Badge size, phenotypic quality, and reproductive success in the house sparrow: a study on honest advertisement. *Evolution* **47**, 1161-1170.

Veiga, J. P. 1995. Honest signalling and the survival cost of badges in the house sparrow. *Evolution* **49**, 570-572.

Verhulst, S. & Tinbergen, J. M. 1991. Experimental evidence for a causal relationship between timing and success of reproduction in the great tit *Parus m. major*. *Journal of Animal Ecology* **60**, 269-282.

Verhulst, S., van Balen, J. H. & Tinbergen, J. M. 1995. Seasonal decline in reproductive success of the great tit: variation in time or quality? *Ecology* **76**, 2392-2403.

von Schantz, T., Göransson, G., Andersson, G., Fröberg, I., Grahn, M., Helgée, A. & Wittzell, H. 1989. Female choice selects for a viability-based male trait in pheasants. *Nature* **337**, 166-169.

Wagner, R. H., Schug, M. D. & Morton, E. S. 1996. Condition-dependent control of paternity by female purple martins: implications for coloniality. *Behavioral Ecology and Sociobiology* **38**, 379-389.

Weatherhead, P. J. & Boag, P. T. 1995. Pair and extra-pair mating success relative to male quality in red-winged blackbirds. *Behavioral Ecology and Sociobiology* **37**, 81-91.

Welling, P., Koivula, K. & Lahti, K. 1995. The dawn chorus is linked with female fertility in the willow tit *Parus montanus*. *Journal of Avian Biology* **26**, 241-246.

West-Eberhard, M. J. 1983. Sexual selection, social competition, and speciation. *Quarterly Review of Biology* **58**, 155-183.

Westerdahl, H., Bensch, S., Hanson, B., Hasselquist, D. & von Schantz, T. 1997. Sex ratio variation among broods of great reed warblers *Acrocephalus arundinaceus*. *Molecular Ecology* **6**, 543-548.

Westneat, D. F. 1990. Genetic parentage in the indigo bunting: a study using DNA fingerprinting. *Behavioral Ecology and Sociobiology* **1990**, 67-76.

Westneat, D. F. & Sherman, P. W. 1997. Density and extra-pair fertilizations in birds: a comparative analysis. *Behavioral Ecology and Sociobiology* **41**, 205-215.

Westneat, D. F. & Webster, M. S. 1994. Molecular analysis of kinship in birds: interesting questions and useful techniques. In: *Molecular Ecology and Evolution: Approaches and Applications.* (Ed. by B. Schierwater, B. Streit, G. P. Wagner & R. DeSalle), pp. 91-126. Basel: Birkhauser Veralg.

Wetton, J. H., Burke, T., Parkin, D. T. & Cairns, E. 1995. Single-locus DNA fingerprinting reveals that male reproductive success increases with age through extra-pair paternity in the house sparrow (*Passer domesticus*). *Proceedings of the Royal Society of London, Series B* **260**, 91-98.

Whitney, C. L. 1981. Patterns of singing in the varied thrush: II. A model of control. *Zeitschrift für Tierpsychologie* **57**, 141-162.

Wildenthal, J. L. 1965. Structure in primary song of the mockingbird (*Mimus polyglottos*). *Auk* **82**, 161-189.

Wiley, R. H. & Poston, J. 1996. Indirect mate choice, competition for mates, and coevolution of the sexes. *Evolution* **50**, 1371-1383.

Wilkinson, G. S., Presgraves, D. C. & Crymes, L. 1998. Male eye span in stalk-eyed flies indicates genetic quality by meiotic drive suppression. *Nature* **391**, 276-279.

Williams, G. C. 1966. *Adaptation and Natural Selection*. Princeton: Princeton University Press.

Williams, G. C. 1979. The question of adaptive sex ratio in outcrossed vertebrates. *Proceedings of the Royal Society of London, Series B* **205**, 567-580.

Wood-Gush, D. G. & Gilbert, A. B. 1970. The rate of egg loss through internal laying. *British Poultry Science* **11**, 161-163.

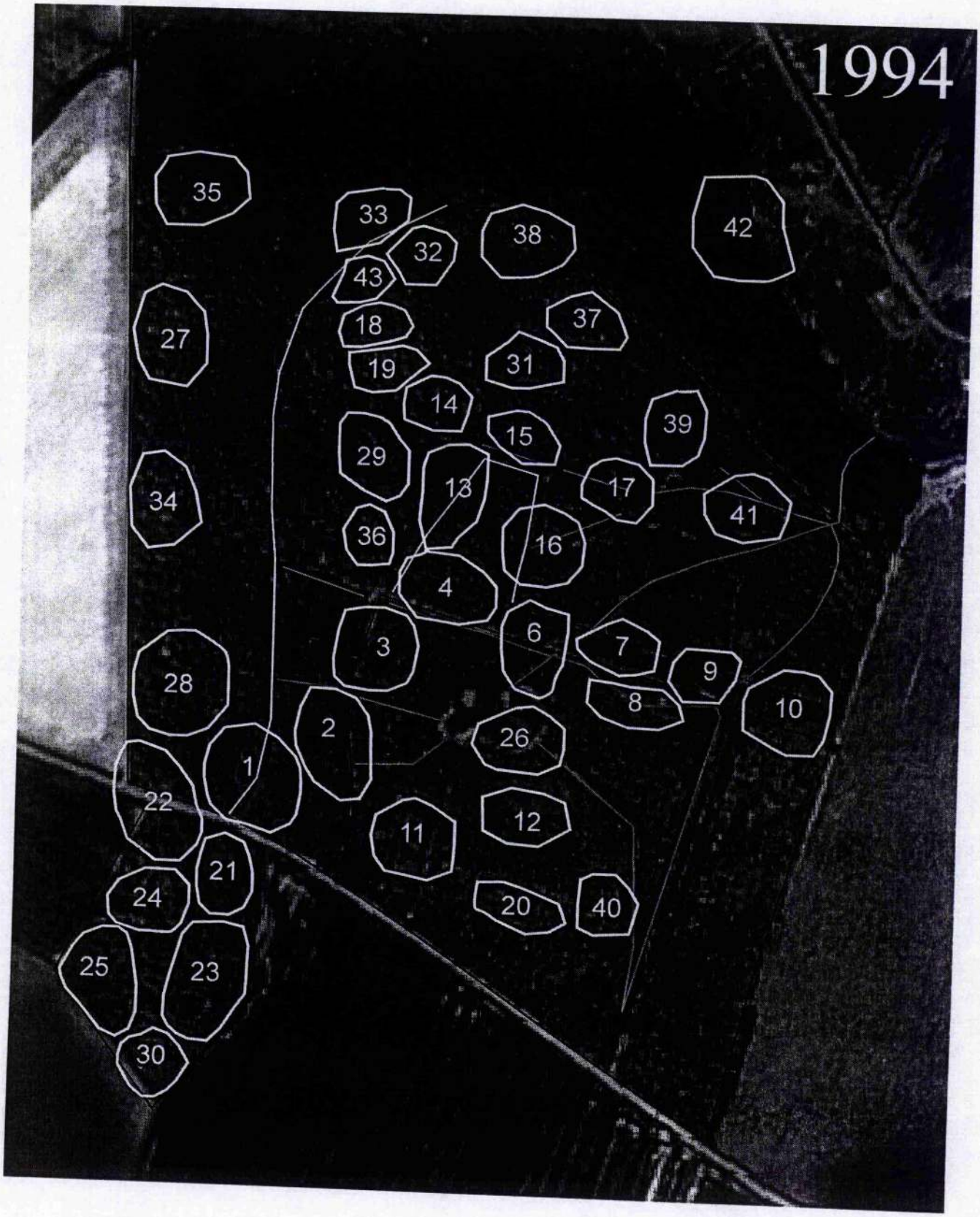
Yasukawa, K. 1981. Song repertoires in the red-winged blackbird (Agelaius phoeniceus): a test of the Beau Geste hypothesis. *Behavioral Ecology and Sociobiology* **29**, 114-125.

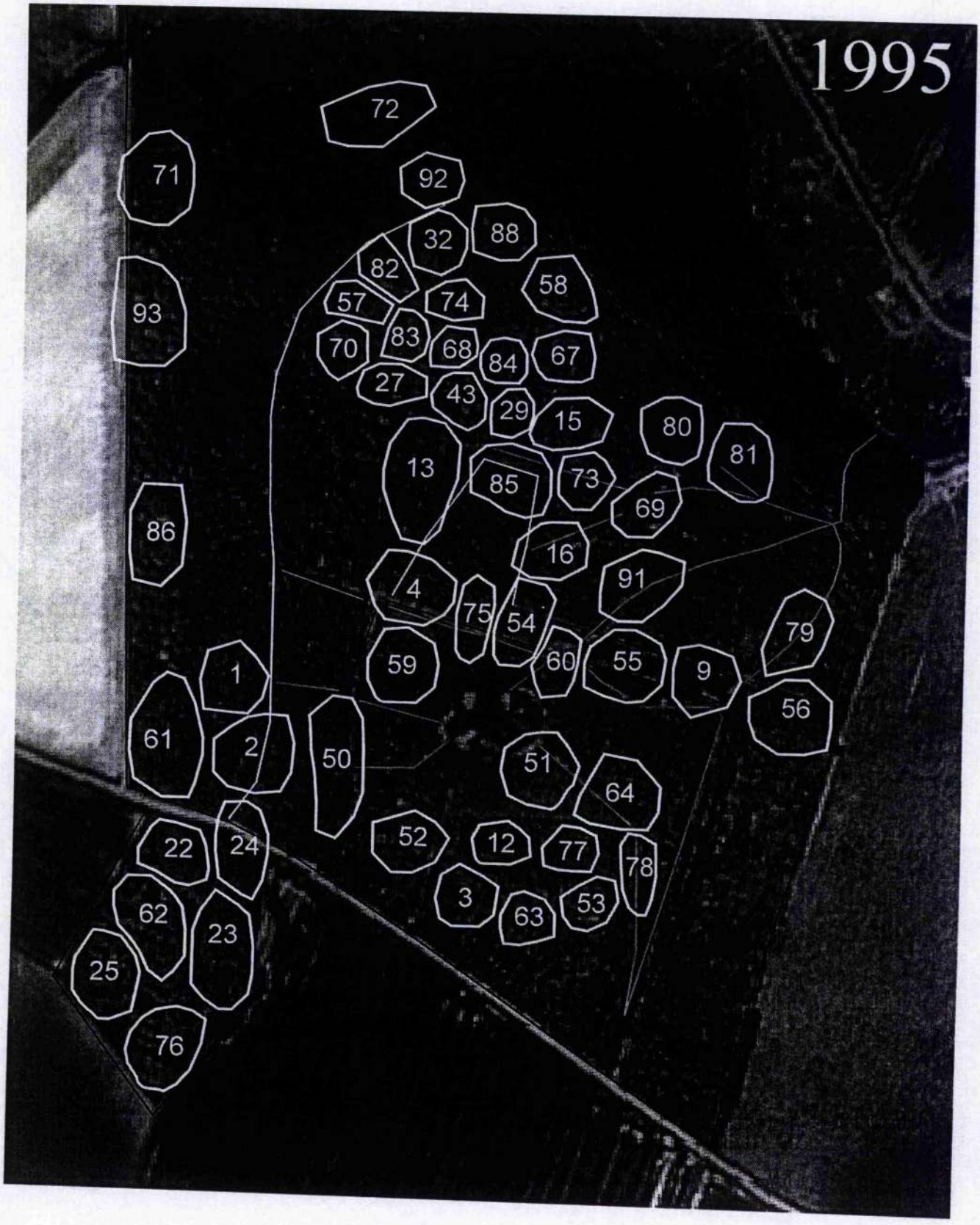
Yasukawa, K., Blank, J. L. & Patterson, C. B. 1980. Song repertoires and sexual selection in the red-winged blackbird. *Behavioral Ecology and Sociobiology* **7**, 233-238.

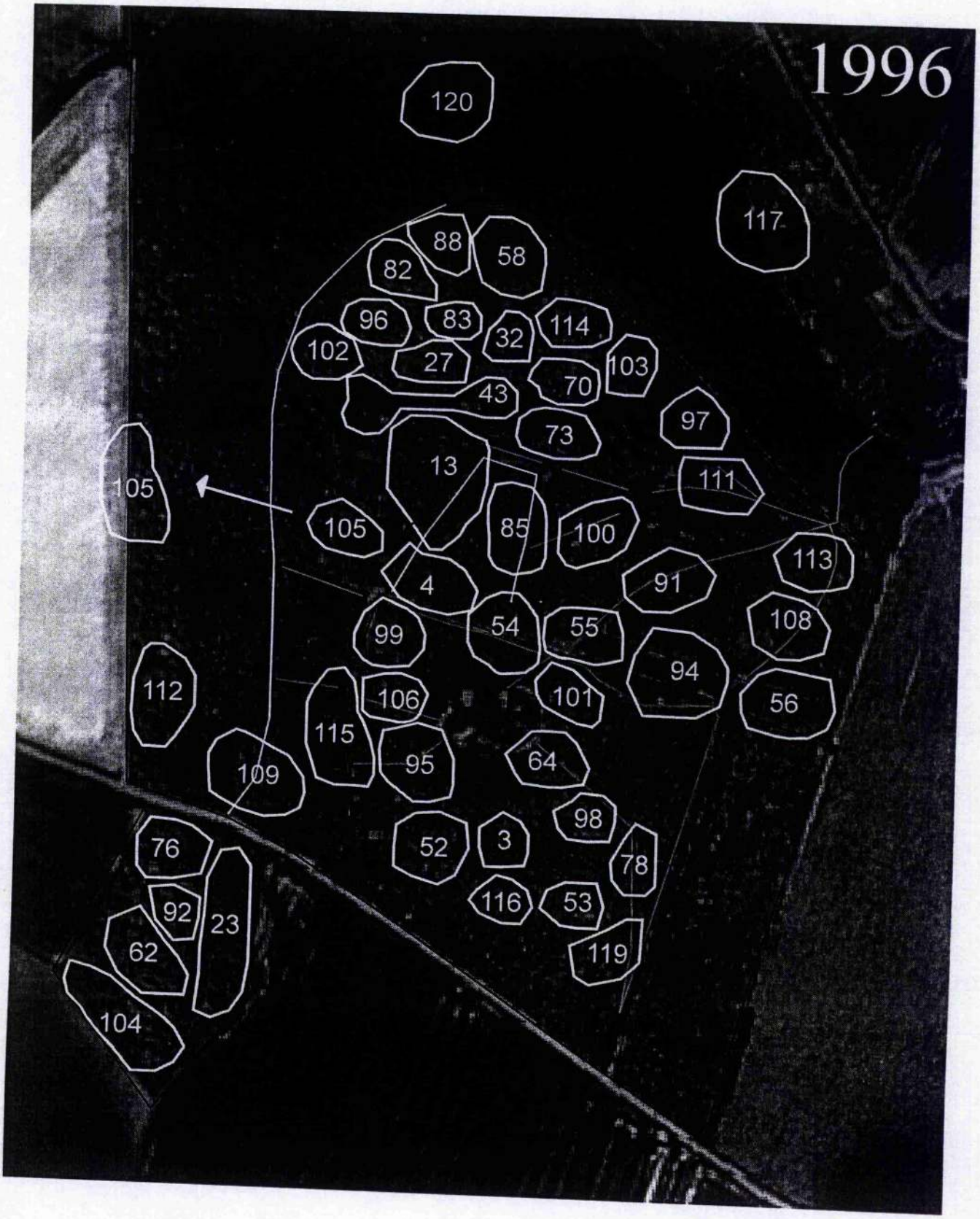
Yezerinac, S. M. & Weatherhead, P. J. 1997. Extra-pair mating, male plumage coloration and sexual selection in yellow warblers (*Dendroica petechia*). *Proceedings of the Royal Society of London, Series B* **264**, 527-532.

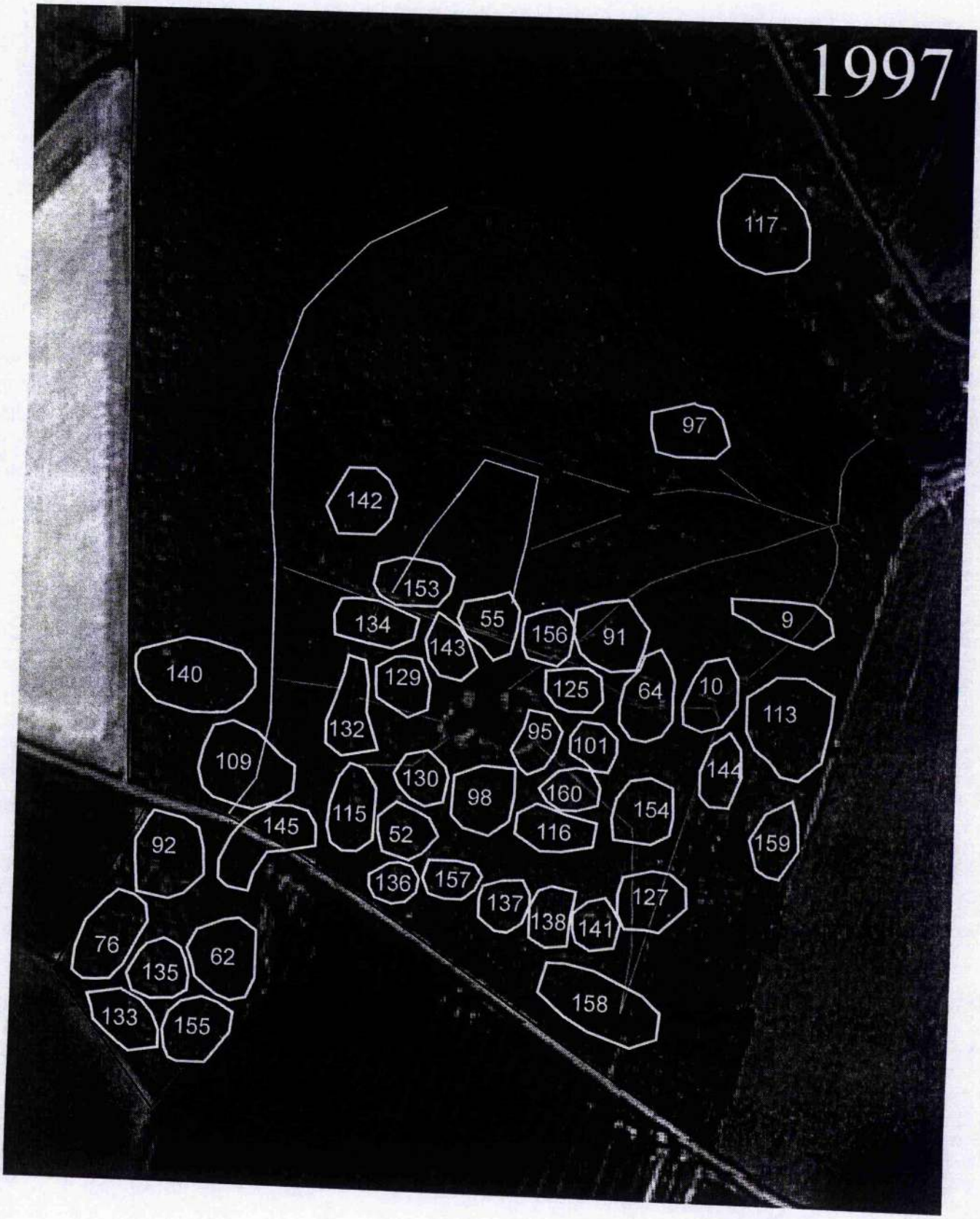
Yezerinak, S. M., Weatherhead, P. J. & Boag, P. T. 1995. Extra-pair paternity and the opportunity for sexual selection in a socially monogamous bird (*Dendroica petechia*). *Behavioral Ecology and Sociobiology* **37**, 179-188.

Zahavi, A. 1975. Mate selection: a selection for a handicap. *Journal of theoretical Biology* **53**, 205-214.


1994
35
33
38
42
32
43
27
18
37
19
31
14
39
15
29
17
34
13
41
36
16
4
3
6
7
9
28
8
10
2
26
1
22
12
11
21
24
20
40
25
23
30


1995
72
71
92
32
88
82
58
57
74
93
70
83
68
84
67
27
43
29
15
80
81
13
85
73
69
86
16
91
4
75
54
79
59
60
55
9
1
56
61
2
50
51
64
52
12
77
78
22
24
3
63
53
62
23
25
76


1996
120
117
88
58
82
96
83
32
114
102
27
103
70
43
97
73
13
111
105
105
85
100
113
4
91
99
54
55
108
94
112
106
101
56
115
95
64
109
98
52
3
78
76
116
53
92
23
62
119
104


1997
117
97
142
153
134
55
156
91
9
143
140
129
125
64
10
113
132
95
101
109
144
130
98
160
154
115
145
52
116
159
92
136
157
137
127
138
141
76
62
135
158
133
155

ERRATA ET ADDENDA

The following typos have been detected:

- page 36, line 17: it should be "another" instead of "other".
- page 78, line 17: it should be "sexes" instead of "sex".
- page 79, the last sentence should be substituted by: "One of these cases is the willow warbler, where variation in song repertoire size does not explain patterns of female choice in the field (Chapter 6)".
- page 93, line 20: it should be "for the variation " instead of "for a the variation".
- Page 127, line 22; it should be "Fig. 5.4" instead of "Fig. 5.5".

The following references have not been duly quoted:

Catchpole, C. K., Dittami, J. & Leisler, B. 1984. Differential responses to male song repertoires in female songbirds implanted with oestradiol. *Nature* **312**, 563-564.

Johnstone, R. A. 1996. Multiple displays in animal communication: 'backup signals' and 'multiple messages'. *Proceedings of the Royal Society of London, Series B* **351**, 329-338.

McGregor, P. K. 1991. Quantifying responses to playback: one, many, or composite multivariate measures? In: *Playback and Studies of Animal Communication.* (Ed. by P. K. MgGregor), pp. 79-96. New York: Plenum Press.

Radesäter, T. S. & Jakobsson, S. 1989. Song rate correlation of replacement territorial willow warblers *Phylloscopus trochilus. Ornis Scandinavica* **20**, 71-73.

Searcy, W. A. & Yasukawa, K. 1996. Song and female choice. In: *Ecology and Evolution of Acoustic Communication in Birds.* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 454-473. Ithaca: Cornell University Press.